DOCUMENT RESUME

ED 181 719 FL 010 854

AUTHOR Njock, Pierre Emmanuel
TITLE La lexicometrie allemande: 1898-1970 (German Lexicometry from 1898 to 1970).
INSTITUTION Laval Univ., Quebec (Quebec). International Center for Research on Bilingualism.
PUB DATE 73
NOTE 126p.
LANGUAGE French

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Computational Linguistics; Dictionaries; *German; *Language Instruction; Language Research; *Lexicography; *Lexicology; Mathematical Linguistics; Research Methodology; *Semantics; Statistics; *Vocabulary; Word Frequency
IDENTIFIERS *Lexicometry

ABSTRACT

The role of lexicometry is to furnish statistical data on all measurable aspects of vocabulary. This study presents an inventory of works on the choice of elements of the German language and outlines the methods of compiling vocabulary with specific reference to the method of choosing words useful for the teaching of a language. It also attempts to isolate problems characteristic of lexicometry in general and German lexicometry in particular, and it addresses the question of a possible evolution in German lexicometry, particulary with regard to criteria for selection. The research covers three stages: (1) a compilation of all works in German lexicometry; (2) an analysis of the available literature from Kaeding (1898) to Pfeffer (1971); and (3) an analysis of the work of Kaeding and Pfeffer. The authors between Kaeding and Pfeffer are divided into two groups, those who have used Kaeding's compilation and those who have not used it. The conclusions are both theoretical and practical. They concern the quality of German words, the method, problems, deficiencies, and evolution of research in lexicometry, the utility of the research, and suggestions for further analyses. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED1817119

centre international de recherches sur le bilinguisme

publication
B-37

LA LEXICOMETRIE ALLEMANDE : 1898 - 1970

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CIRB - Ilam

Langues et Linguistique

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PIERRE EMMANUEL NJOCK

1973

FL010854

CIRB
ICRB

Pierre Emmanuel Njock
Assistant d'allemand
Faculté des Lettres et Sciences Humaines
Université de Yaoundé
Cameroun

LA LEXICOMETRIE ALLEMANDE: 1898 - 1970

Publication B-37

1973
Centre international de recherches sur le bilinguisme
International Center for Research on Bilingualism
Quebec

Le Centre international de recherches sur le bilinguisme est un organisme de recherche universitaire qui reçoit des subventions de soutien du Ministère de l'éducation du Québec et du Secrétariat d'Etat du Canada.

© 1973 CENTRE INTERNATIONAL DE RECHERCHES SUR LE BILINGUISME
Tous droits réservés. Imprimé au Canada
Dépôt légal (Québec): 1er trimestre 1973

RECONNAISSANCE

Notre reconnaissance s'adresse avant tout aux professeurs W.F. Mackey et L. Laforge. Le premier pour avoir su nous intéresser et nous initier à la lexicométrie, le second non seulement pour avoir accepté la direction de cette recherche mais aussi pour le dévouement avec lequel il a tout mis en oeuvre pour que ce travail aboutisse à sa fin. Le professeur Laforge s'est donné la peine de lire la version manuscrite et de faire des suggestions constructives. Les fautes qui subsisteraient ne devraient être imputées cependant qu'au seul auteur.

Notre reconnaissance va également aux autorités du CIRB de l'Université Laval pour leur soutien moral et matériel. Il en est de même de nos amis mais tout particulièrement *Frdulein* Gertrud Deimling, *Pfarrer in Ruhestand* Herbert Fuchs et *Missionar* Walter Trobisch qui nous ont facilité un voyage d'information dans différents centres de recherches en Allemagne.

Au professeur J.A. Pfeffer qui a mis la bibliothèque de son *Institute for Basic German* à notre disposition.

Au personnel de l'Institut Goethe et des autres centres de recherches à Hambourg et à Mannheim notamment.

A tous nos professeurs de didactique des langues et d'allemand.

A nos parents.

Québec, Mars 1973

PEN

TABLE DES MATIERES

RECONNAISSANCE

TABLE DES MATIERES

0. INTRODUCTION 1

0.1 Objectifs 2
0.2 Etapes du travail 2
0.3 La lexicométrie: définition 3

1. CHAPITRE I 7

1.1 Justification du choix de la période 8
1.2 Examen critique de quelques travaux 9
1.3 Importance du travail 10
1.4 Grille d'analyse 10

2. CHAPITRE II 21

2.1 Kaeding, Friedrich Wilhelm 23
2.2 Les travaux basés sur Kaeding 28
2.3 Autres travaux 40
2.4 Pfeffer, J.A. 57
2.5 Table chronologique de la lexicométrie allemande 63

3. CONCLUSIONS 71

3.1 Conclusions théoriques 72
3.2 Conclusions pratiques 74
3.3 Déficiences de l'analyse 74
3.4 Suggestions pour des analyses ultérieures 75

4. BIBLIOGRAPHIE 77

4.1 Abréviations et symboles 77
4.2 Ouvrages 79

5. APPENDICES 98

5.1 Appendice A: Tableau des langues traitées ou en cours de traitement lexicométrique 100
5.2 Appendice B: Principaux centres de traitement 103
5.3 Appendice C: Petit lexique de la lexicométrie 107
5.4 Appendice D: Calculs et notes sur le critère de l'usage 110
5.5 Appendice E: Index des noms d'auteurs parus dans le texte 112
5.6 Appendice F: Résumé du travail 116

INTRODUCTION

0.1 Objectifs

- 0.1.1 Inventaire des travaux allemands
- 0.1.2 Recherche des méthodes de compilation
- 0.1.3 Problèmes dans ce genre de travail
- 0.1.4 Evolution de cette recherche pour la langue allemande

0.2 Etapes du travail

- 0.2.1 Recherche des listes de vocabulaire allemand
- 0.2.2 Analyse
- 0.2.3 Classement

0.3 La lexicométrie: définition

- 0.3.1 Situation
- 0.3.2 Terminologie

0.1 OBJECTIFS

0.1.1 Inventaire des travaux allemands

La présente étude voudrait, avant tout, présenter un inventaire des travaux consacrés au choix des éléments de la langue allemande. En d'autres termes il s'agit de répondre à deux questions: existe-t-il un nombre important de travaux en lexicométrie allemande? Si oui lesquels?

0.1.2 Recherche des méthodes de compilation

Nous voulons également dégager les méthodes de compilation de vocabulaire. Nous nous demanderons ici comment on fait le choix des mots utiles pour l'enseignement d'une langue en général.

0.1.3 Problèmes dans ce genre de travail

Toute recherche présente un certain nombre de problèmes; aussi essayerons-nous de relever ceux qui nous semblent caractériser la lexicométrie en général, et la lexicométrie allemande en particulier. En d'autres mots nous chercherons à répondre à la question suivante: à quels genres de problèmes doit-on s'attendre quand on fait une liste de mots utiles?

0.1.4 Evolution de cette recherche pour la langue allemande

L'examen des différents travaux nous permettra de dire si oui ou non on peut parler d'une évolution en lexicométrie allemande. Il s'agira de voir en quoi les travaux récents diffèrent des tout premiers, surtout au niveau des critères de sélection.

0.2 ETAPES DU TRAVAIL

0.2.1 Recherche des listes de vocabulaire allemand

Cette recherche s'est faite en trois grandes étapes. La première, longue et pénible, consistait à déterminer et à compiler les travaux de lexicométrie allemande. Elle était d'autant plus difficile que le terme lexicométrie n'était ni connu ni clairement défini avant 1965. Il fallait compléter les quelques titres découverts dans les bibliographies des ouvrages de statistique linguistique avec ceux obtenus soit par correspondance, soit par contact direct avec les centres de recherches aux Etats-Unis et en Allemagne.

0.2.2 Analyse

Ayant relevé un certain nombre de titres, nous avons commencé à en faire l'analyse par ordre chronologique et d'après une grille que nous avons conçue (1.1.5). L'accent a été mis sur les travaux exécutés dans la période qui s'étend de Kaeding (1898) à Pfeffer (1971).

0.2.3 Classement

A chacun de ces deux auteurs nous avons consacré un chapitre. Quant aux autres, nous les avons classés en deux grands groupes. Le premier réunit les travaux qui ont utilisé les résultats de la compilation de Kaeding; il représente presque la moitié des travaux analysés. L'autre groupe inclut les auteurs dont les travaux n'utilisent pas la liste de Kaeding.

0.3 LA LEXICOMETRIE: DEFINITION

0.3.1 Situation

L'objet de notre recherche se rattache à la didactique des langues et particulièrement à la didactique analytique. Cette dernière comprend entre autres l'analyse des méthodes, l'utilisation des méthodes et leurs résultats dans l'apprentissage des langues. Le premier de ces secteurs inclut à son tour, l'analyse du choix, de l'ordre, de la présentation et de la répétition des éléments. La lexicométrie s'intéresse avant tout à la sélection des éléments.

0.3.2 Terminologie

Afin de mieux préciser le sens du terme lexicométrie, nous allons d'abord procéder à un certain nombre de rappels. Une langue est essentiellement constituée de sons, de mots et de constructions. Ceci a conduit à la distinction de trois disciplines linguistiques: l'étude des sons ou phonèmes, la phonétique, l'étude des mots ou lexicologie et l'étude des constructions ou syntaxe. La lexicométrie se rapproche plutôt de la lexicologie. Cette dernière étudie les mots d'une langue soit dans leur formation, soit dans leur signification et ce, à trois points de vue différents: le point de vue descriptif, la morphologie, étudie les procédés de formation des mots (suffixation, préfixation, composition); la morphologie étudie les mots en tant qu'éléments flexionnels, c'est-à-dire capables de revêtir une forme déterminée selon leur fonction. Le point de vue historique, l'étymologie, étudie les mots dans leur origine, c'est-à-dire les causes qui ont déterminé la naissance des mots, leur usage, leur abandon, leur concurrence avec d'autres mots. Le point de vue normatif fait l'étude des relations dans le vocabulaire d'une

langue. Si la lexicométrie s'intéresse aussi au vocabulaire on évitera cependant de l'assimiler à la lexicologie, en effet elle ajoute un autre aspect dans son étude; qui dit sélection suppose un choix entre un certain nombre, et qui parle de nombre entre dans le domaine des mathématiques et notamment des statistiques. La spécialité qui fait l'objet de cette étude a alors porté divers noms: En 1949 Michéa (142) parle "d'une science encore jeune, mais riche d'avenir, la lexicologie statistique". Cette appellation fait justement ressortir les liens de la lexicométrie avec d'une part la lexicologie et d'autre part les statistiques. Quinze ans plus tard Juilland (101, p. XLVII) parle aussi de *statistical lexicographer* pour désigner les spécialistes de la sélection du vocabulaire. Mais on rencontre aussi d'autres termes comme linguistique mathématique, statistique linguistique (167) et linguistique quantitative (71). Certains allemands ont utilisé les mots *die mathematisch-statistische Methode in der Sprachwissenschaft* (83, p. 47). Cette définition nous paraît très longue. Hartmann (83) lui-même utilise *Worthäufigkeitszählung (-forschung)*, terme qui se comprend puisque la majorité des travaux faits jusqu'à date se fondent sinon exclusivement du moins principalement sur l'étude de la fréquence. Maintenant qu'on sait qu'il y a d'autres critères objectifs de sélection, on devrait alors chercher un terme plus convenable à la spécialité qui fait l'objet de la présente étude. Et c'est chez un professeur d'Euskirchen (République Fédérale d'Allemagne) que nous croyons l'avoir trouvé. Il s'agit de Rolf Dietrich Keil (104) qui dit en l'occurrence:

> *"Lexikometrie (= Wortschatzmessung) wird von mir, im Anschluss an eine Anregung von de Tollenaere, als Terminus vorgeschlagen für jenen Zweig der Lexikologie (= Wortschatzkunde), der mit statistischen Methoden Daten für alle messbaren Erscheinungen des Wortschatzes liefert."*

Cette définition présente la lexicométrie non pas comme une science indépendante, mais plutôt comme une science auxiliaire *Hilfswissenschaft*, une branche de la lexicologie, dont le rôle est de fournir des données statistiques sur tous les aspects mesurables du vocabulaire. Le terme *lexicométrie* est encore jeune. Il n'apparaît dans aucun des grands dictionnaires de l'heure, même pas dans celui de Marouzeau (125) ni de Nash (171. a). Nous ne l'avons rencontré nulle part dans un écrit publié avant 1965. Pour résumer nos définitions on pourrait dire que la lexicographie est la science des dictionnaires, c'est-à-dire la technique de leur fabrication, la lexicologie est l'étude du vocabulaire et la lexicométrie la mesure du vocabulaire. Cette dernière branche dessert les deux précédentes mais pas exclusivement; elle a aussi d'autres fins.

Maintenant qu'on a une définition de lexicométrie on pourrait se demander à quoi cette spécialité sert. Nous n'allons pas nous étendre sur la réponse à cette question. Nous renvoyons plutôt le

lecteur intéressé aux travaux de Keil (104), Guiraud (71, p. 7 ff), Keller (105), Price, W. (196), Laforge (112). Nous aimerions cependant attirer l'attention sur certains points: les indices de fréquence des mots, les index d'auteurs ou de texte, et les thesaurus établis par les techniques lexicométriques peuvent fournir entre autres, des renseignements sur les morphèmes; les données ainsi obtenues sont utiles pour la transmission du vocabulaire fondamental et du vocabulaire minimum spécialisé au cours élémentaire. Elles sont aussi utiles pour l'analyse littéraire, on peut comparer les auteurs, les styles et les époques. Ces mêmes renseignements sur les morphèmes sont utiles pour l'introduction des formes morphologiques, d'après leur fréquence. Ces données aideront par ailleurs à trouver la nature et la fonction des mots ainsi que leur participation dans la totalité du vocabulaire des langues non standardisées ou en voie de standardisation. Les techniques lexicométriques peuvent être d'un grand secours au grammairien de ces langues, notamment dans la détermination de la nature et de la fonction des mots.

CHAPITRE I

FONDEMENTS DE LA RECHERCHE

1.1 Justification du choix de la période

1.2 Examen critique de quelques travaux

1.3 Importance du travail

1.4 Grille d'analyse

- 1.4.1 Buts d'une compilation
- 1.4.2 Critères dominants de sélection
- 1.4.3 Corpus
- 1.4.4 Méthode
- 1.4.5 Problèmes spécifiques
- 1.4.6 Accessibilité
- 1.4.7 Innovations majeures
- 1.4.8 Utilisations ultérieures

Le lecteur curieux se posera probablement de nombreuses questions dont: pourquoi la période de Kaeding à Pfeffer a-t-elle été choisie et non pas une autre? N'existe-t-il pas déjà beaucoup de travaux sur ce sujet? En quoi ce travail diffère-t-il des autres? A qui et à quoi pourra-t-il servir? Comment l'auteur entend-il faire son analyse? Essayer de répondre à chacune de ces questions est l'objet du présent chapitre.

1.1 JUSTIFICATION DU CHOIX DE LA PERIODE

Si le terme *lexicométrie* est jeune, il n'en est pas de même du genre de travail qu'il représente. En 1869 Meyer (139) publie à Berlin un travail de décompte de la langue parlée; sept ans plus tard paraît dans la même ville, un travail sur la concordance. Kaeding (102) signale aussi de nombreuses autres études entreprises et parfois publiées avant son enquête. Ces études se rapportent soit à la fréquence des mots, soit à la fréquence des sons. Mais elles ne reposent que sur un corpus très peu important, 100 000 syllabes environ, et ne cherchent à répondre qu'à des questions spéciales comme l'abréviation des noms concrets ou la fréquence des auxiliaires. Il faudra attendre la fin du dix-neuvième siècle pour voir naître un travail d'envergure.

F. de Tollenaere (241, p. 110) faisait une fois remarquer qu'au XIXème siècle, c'est en Allemagne que s'allume le flambeau lexicographique. On pourrait dire la même chose de la lexicométrie; en effet un coup d'oeil sur les travaux de lexicométrie montre que le plus grand travail jamais réalisé - surtout à la main - est celui d'un allemand vers la fin du XIXème siècle: il s'agit de l'oeuvre monumentale de Kaeding. C'est elle qui semble ouvrir la porte à d'autres chercheurs en leur fournissant une méthode de compilation. Plusieurs listes de vocabulaire voient le jour quelques années après la recherche de Kaeding. Notre curiosité nous a donc poussé à tourner les yeux vers les allemands et précisément à partir de la fin du XIXème siècle. On compte plus d'une quarantaine de lexicométristes-germanistes depuis le début du XXème siècle. L'un des contemporains, le plus remarquable sans doute par sa méthode de travail et ses publications, est Pfeffer (2.4). Son *Institute for Basic German* en était en 1970 à sa septième publication et travaillait activement sur d'autres projets sur l'allemand fondamental parlé et écrit.

L'examen de la période allant de Kaeding à Pfeffer pourra donc nous montrer assez clairement la situation de la lexicométrie allemande ainsi que son évolution pendant environ trois quarts de siècle. En d'autres termes, l'étude de cette période est suffisante pour atteindre les objectifs que nous avons posés au début de cette recherche.

1.2 EXAMEN CRITIQUE DE QUELQUES TRAVAUX

L'une des tâches les plus difficiles dans cette étude était l'établissement de la documentation. Ceci provient du fait que nous n'avions pas encore vu ni entendu parler d'un travail semblable au nôtre. Bon nombre d'ouvrages nous ont cependant été d'une grande utilité. Nous n'avons pas l'intention de les citer tous ici, nous mentionnons cependant le *Language Teaching Analysis* de Mackey dont une section de la bibliographie (119, Nos 1479-1496) est réservée aux ouvrages de sélection du vocabulaire allemand. L'article de Keil (104) a été notre tremplin. C'est à lui que nous avons emprunté la définition du terme lexicométrie. Keil évoque, dans son article, l'importance scientifique et pratique de la lexicométrie. Il explique pourquoi cette importance ne se fait pas encore sentir après tant de travaux pour enfin exposer, en guise d'introduction à une discussion du problème, ce qui lui paraît être le déroulement convenable des opérations dans un travail de compilation. Il donne aussi une bibliographie importante de la lexicométrie.

L'ouvrage de F. de Tollenaere (241) nous a fourni bon nombre d'indications sur les différents centres de lexicométrie en Europe.

Hartmann (83) a été à la base de notre tableau chronologique des travaux lexicométriques. Il montre l'importance de l'étude de la fréquence des mots dans divers domaines. Il présente aussi l'évolution historique de cette recherche en groupant le matériel bibliographique en tableaux.

La bibliographie de Coleman (21, 22, 23) nous a été d'un grand secours. Elle couvre la période de 1927-1942 en trois volumes. Elle est assez exhaustive quant aux limites de ses intérêts, notamment pour les publications aux Etats-Unis et au Canada. Bien qu'incluant certaines contributions importantes d'Europe, elle est loin de nous donner suffisamment de détails sur les travaux d'Europe en général et d'Allemagne en particulier.

Le livre de Savard et Richards (211) offre un aperçu historique de la lexicométrie. Il résume les contributions de Leibniz et de Kaeding puis signale les recherches de Morgan, Bakonyi, Hauch et Schneider. Mais l'accent porte plus sur les vocabulaires de base du français et de l'anglais.

Nous ne pouvons passer sous silence l'ouvrage de Laforge, *La sélection en didactique analytique* (112) d'où nous avons emprunté bon nombre de définitions et dont l'importante bibliographie nous a été utile.

La bibliographie internationale de Nostrand (175) couvre la période de 1945-1961. Il y manque donc neuf ans (1962-1971) assez fructueux pour la lexicométrie allemande.

De tous les écrits qui nous ont servi comme source primaire ou secondaire aucun n'a cependant pour but de présenter la lexicométrie allemande seule; même pas les articles les plus spécialisés en statistique linguistique. Les différents auteurs se sont intéressés jusqu'à présent plutôt à ce que nous appellerions la lexicométrie générale qu'à la lexicométrie spécialisée. Par *générale* nous entendons ce qui englobe tous les travaux des différentes langues et par *spécialisée* la lexicométrie se limitant à une seule langue comme l'anglais, l'allemand ou le français.

Le tableau chronologique de Helmut Meier (131, p. 349-351) constitue une légère exception à ce qui vient d'être dit, mais il ne couvre que la période allant jusqu'à 1963 et l'auteur ne visait pas l'exhaustivité. La majorité de ses titres indique des travaux de statistique linguistique allemande. Comme il s'agit alors d'une table chronologique, les renseignements fournis par l'auteur ne sont pas suffisants pour un inventaire bibliographique.

1.3 IMPORTANCE DU TRAVAIL

On peut se demander s'il vaut la peine de préparer un inventaire historique de la lexicométrie allemande. Cette question se pose surtout lorsqu'on consulte les principales bibliographies de la statistique linguistique. En effet aucune de ces bibliographies n'inclut un nombre assez grand de travaux allemands et ce qui frappe alors c'est la disproportion entre la quantité de compilations allemandes et celle des compilations d'autres langues réunies. Mais un examen poussé montre qu'il existe au moins une quarantaine de compilations pour la seule langue allemande, ce qui, pour une spécialité encore jeune comme la lexicométrie, est tout de même significatif, et nous ne prétendons pas que notre bibliographie soit déjà exhaustive.

1.4 GRILLE D'ANALYSE

Afin d'analyser plus systématiquement les différents travaux, il nous a fallu concevoir une grille. Celle-ci devant nous permettre de regrouper les lexicométristes d'après ce qu'ils peuvent avoir de commun. Il s'agit donc, dans ce paragraphe, d'exposer les procédés et les techniques que nous comptons mettre en oeuvre pour cueillir nos données.

Il existe à l'heure actuelle une grande divergence dans les travaux de statistique du vocabulaire - cette divergence est une des raisons pour lesquelles on considère difficilement la lexicométrie comme une science. Nous espérons cependant que cette branche de la lexicologie ne tardera pas à s'ériger en science - Keil (104) invite tous les lexicométristes à y réfléchir. Les listes diffèrent généralement par les buts que vise l'auteur de la compilation,

les critères de sélection, le corpus ou population lexicale, la définition de l'unité de décompte, la présentation des résultats. Nous examinerons donc chacun de ces points. Pour répondre aux objectifs posés dans l'introduction, nous avons également prévu de faire ressortir les innovations majeures des auteurs, ainsi que les utilisations dont leurs travaux ont pu être l'objet.

1.4.1 Buts d'une compilation

Une des premières questions qu'on peut se poser en étudiant les travaux de sélection de vocabulaire est de savoir le but que cherche l'auteur. Son vocabulaire est-il pour l'enseignement traditionnel? pour les voyages? pour des tests? On pourrait étendre la question en se demandant à qui se destine ce vocabulaire: est-il préparé à l'intention des enfants, des adolescents ou des adultes, à l'intention des étrangers ou des autochtones? La présente étude s'intéressant spécialement à la didactique des langues, nous avons choisi de classer les travaux de lexicométrie en deux groupes d'après les buts de l'auteur: l'enseignement des langues et autres buts.

1.4.1.1 Enseignement des langues

La plupart des travaux visent l'enseignement traditionnel des langues. Cet enseignement peut être conçu pour:

a) une langue maternelle

C'est le cas du *Recuento de Vocabulario español* (204). Ce travail est basé sur la notion de fréquence (1.4.2.1) et vise surtout l'enseignement de l'espagnol comme langue maternelle. Il essaie de refléter aussi bien la langue écrite que la langue parlée de différents âges et niveaux sociaux dans toutes les régions de l'île de Puerto Rico. Thorndike (239) a préparé également sa liste de mots pour l'enseignement de la langue maternelle. Il vise la lecture des textes. Quant à Horn (98) c'est l'enseignement de l'écriture qu'il recherche.

b) une langue seconde

C'est dans cette catégorie que tombe la majorité des recherches. Les vocabulaires de base ont été dès le début un moyen de faciliter l'apprentissage de la langue à un étranger. Le *français fondamental* (155), le *General Service List of English Words* (265) ont contribué à l'enseignement et à l'apprentissage du français et de l'anglais comme langue seconde.

1.4.1.2 Autres buts

Sous cette rubrique nous inclurons les travaux ne visant pas l'enseignement classique de la langue. C'est comme le travail de Kaeding (2.1) dont le but est d'améliorer le système de sténographie.

1.4.2 Critères dominants de sélection

La sélection est le premier des quatre principes qui caractérisent une méthode d'enseignement. Voici ce qu'en dit Mackey (119, 157):

> *"All teaching, whether good or bad, must include some sort of selection, some sort of gradation, some sort of presentation, and some sort of repetition. Selection, because it is impossible to teach the whole of a field of knowledge; we are forced to select the part of it we wish to teach."*

La sélection apparaît donc comme un principe obligatoire. Qu'on fasse l'apprentissage et/ou l'enseignement d'une langue, on est bien obligé de faire le choix des matériaux utiles. Les lexicométristes disposent de plusieurs critères de sélection. Certains chercheurs n'en utilisent qu'un seul dans leurs travaux, d'autres deux ou plus. Savard et Richards (211) en ont proposé quatre. La fréquence, la répartition, l'usage, la valence, la disponibilité, la familiarité, la facilité et l'empirisme sont certains de ces critères que nous voulons rappeler dans les lignes qui suivent.

1.4.2.1 La fréquence

La fréquence est probablement le critère le plus utilisé dans la sélection des matériaux d'enseignement. Elle est un caractère des éléments du discours. Laforge (112, p. 163) la définit comme "... le nombre d'occurrences d'un élément dans une portion ou un échantillon de discours, le véhicule de cet échantillon étant soit l'enregistrement de la langue parlée, soit un texte écrit." Certains chercheurs estiment qu'un élément très fréquent dans un discours est généralement utile dans l'apprentissage de la langue. La fréquence comme indice d'utilité d'un élément varie selon la nature et la grandeur de l'échantillon. En effet on s'est rendu compte qu'un mot pouvait être fréquent dans un texte donné ou dans une partie d'un texte, sans qu'il le soit ailleurs. Les noms concrets sont ceux qui tombent le plus souvent dans cette catégorie de mots. Ceci fausse alors un peu la notion de fréquence par l'importance qu'on peut accorder à un mot de ce genre. Les chercheurs ont donc eu recours à un critère complémentaire.

1.4.2.2 La répartition

La répartition sert en quelque sorte à compléter les informations apportées par l'indice de fréquence. En effet elle ajoute la notion de stabilité de fréquence, autrement dit de sa distribution dans un corpus. En lexicométrie on assume qu'un élément que l'on retrouve partout sera plus important que celui qu'on ne retrouve que dans un seul texte, même si sa fréquence est très élevée. La répartition indique le nombre de sous-ensembles dans lesquels une unité de décompte est représentée. Vander Beke et Thorndike l'ont appelée *range* et c'est Gougenheim qui a proposé le terme *répartition*. Si certains travaux comme celui de Vander Beke accordent une plus grande importance à la répartition, il faut cependant constater que ce critère n'a pas encore été utilisé seul dans une sélection de vocabulaire, mais plutôt comme un des critères objectifs accompagnant ou accompagné de la fréquence et/ou d'autres critères (voir infra 1.4.2.9). Juilland (101) l'appellerait *simple dispersion*. C'est de ce chercheur que nous tenons le critère suivant.

1.4.2.3 L'usage

La notion d'usage résulte de celles de fréquence et de répartition. Elle est née d'une méthode mise au point grâce au dépouillement mécanographique des textes et à l'usage des ordinateurs. Cette méthode est présentée et appliquée par Juilland de l'université de Stanford dans *Frequency Dictionary of Spanish Words* (101).

Si les sous-ensembles d'une population lexicale sont égaux on peut quantifier la dispersion et l'utiliser pour l'estimation de l'usage du mot - nous adoptons la traduction *usage* de *word usage* que Muller donne dans son article (166). Il existe deux techniques de base pour calculer l'usage des mots. Elles tiennent compte de la dispersion. L'une d'elle est de considérer la dispersion comme le nombre de sous-ensembles dans lesquels le mot apparaît. Juilland l'appelle *simple dispersion* et ce n'est rien d'autre que la répartition. Plusieurs auteurs l'ont utilisée comme critère de sélection pour leurs vocabulaires de base: c'est ainsi que Keniston (106) divise son vocabulaire espagnol en six classes de dispersion. Il en est de même de Garciá Hoz (55) dont la liste contient tous les mots se retrouvant dans les quatre sous-ensembles de sa population, quelle que soit leur fréquence. Mais on s'est rendu compte que la *simple dispersion* ne suffit pas. En effet des mots de même fréquence et de même répartition mais mal représentés dans leurs sous-ensembles peuvent être mal évalués. Nous illustrons ce fait par un exemple: soit deux mots A et B de même fréquence f = 210 et de même répartition (ou simple dispersion) r = 7; leur fréquentation des sept sous-ensembles étant la suivante:

A -	1	1	1	1	1	1	204
B -	30	30	30	30	30	30	30 [1]

A fréquence et répartition égales cet exemple montre qu'il y a une différence entre les deux mots. B est bien plus régulier, donc mieux réparti et par conséquent plus utile que le mot A qui, sans être inutile, doit plutôt être pris comme un mot disponible dans le dernier sous-ensemble. Les lexicométristes ont alors cherché des techniques plus complexes pour remédier à cette lacune de la répartition. Cette recherche les a conduits à la deuxième technique qui tient compte de ce que Juilland appelle *complex dispersion* - nous disons seulement dispersion -. Cette technique a recours à des formules qui mesurent la régularité et l'uniformité de la répartition des occurrences du mot dans les sous-ensembles d'une population lexicale. Elle repose non seulement sur le nombre de sous-ensembles où le mot se trouve, mais aussi sur sa contribution effective dans chacun d'eux. Juilland trouve une formule quantifiant la dispersion de tout mot de la population par un coefficient qui varie entre 0 (dispersion minimale) et 1 (dispersion maximale). Il s'agit d'une formule qui, après calculs se réduit à

$$D = 1 - \frac{\sqrt{}}{\sqrt{n-1}} \quad \text{(166, p. 37) où}$$

$\sqrt{} = \frac{\sigma}{\bar{x}}$ = coefficient de variation;
n = nombre de sous-ensembles;
σ = écart-type;
$\bar{x}$ = moyenne théorique.

L'indice d'usage s'obtient alors en faisant le produit de la fréquence et de la dispersion.

$$U = F \times D \text{ où}$$

U = coefficient d'usage
F = fréquence
D = indice de dispersion

Cette formule de l'usage a l'avantage de posséder un pouvoir de discrimination remarquable. Elle assigne différents coefficients d'usage aux mots de même fréquence et de même répartition, mais dont les coefficients de dispersion sont différents.

---

1. Voir calculs infra p. 110-111.

1.4.2.4 La disponibilité

Les enquêtes basées sur la fréquence et la répartition font apparaître des lacunes quant à la stabilité des noms concrets. Michéa s'étant rendu compte de ce fait groupe les noms concrets par centres d'intérêt afin de mieux mesurer leur importance. C'est de Michéa qu'on a la définition suivante:

> "un mot disponible est un mot qui, sans être particulièrement fréquent, est cependant toujours prêt à être employé, et se présente immédiatement à l'esprit au moment où l'on en a besoin" (149).

Il s'agit essentiellement d'un mot thématique, c'est-à-dire celui qui exprime la chose en elle-même. Les mots thématiques s'opposent aux mots athématiques,

> "ces mots qui servent à nous exprimer au sujet des choses plutôt qu'à exprimer les choses elles-mêmes, des termes plus ou moins communs à tous les sujets, à toutes les situations" (123, p. 23).

Si la fréquence, la répartition et l'usage permettent de trouver les mots utiles, il est vrai que le vocabulaire basé sur ces critères reste dominé par des mots-outils et des mots abstraits. L'utilité des mots concrets ne se mesure que par la disponibilité. La première enquête sur le vocabulaire disponible est celle de Michéa. Elle avait été conçue pour l'élaboration d'un vocabulaire fondamental du français comme langue seconde. Une deuxième enquête importante a paru chez Didier en 1971, il s'agit du *Vocabulaire disponible du français*, en deux tomes, publié par Mackey, Savard et Ardouin. C'est une étude témoin qui doit servir de base aux lexicométristes, notamment à ceux intéressés à l'étude du critère de la disponibilité. Elle apporte une bonne documentation pour la comparaison du vocabulaire commun de la France et de l'Acadie. C'est la première étude qui utilise la disponibilité comme mesure des différences culturelles et établit des indices de disponibilité à l'aide de l'ordinateur électronique. Toutefois Mackey précise dans la préface que

> "l'étude ne va pas assez loin pour établir les comparaisons entre les deux vocabulaires. Elle se borne à fournir les outils à ceux qui entreprennent une telle comparaison. Elle n'exploite pas la documentation, elle explique simplement comment cette documentation a été obtenue."

Une enquête sur la disponibilité reste souvent très limitée parce qu'elle ne porte que sur un certain nombre de centres d'intérêt. Elle ne touche donc que quelques noms concrets, quelques verbes et quelques adjectifs des centres choisis. Les lexicométristes

ne se sont pas arrêtés qu'à la disponibilité, ils ont trouvé un autre critère.

1.4.2.5 La valence lexicale

Nous donnons au terme valence une acception différente de celle des maîtres allemands contemporains. Leur terme *Valenz* qui est présentement en vogue, notamment à l'*Institut für deutsche Sprache* à Mannheim et à Freiburg, mais aussi à Leipzig, se limite pour l'instant aux verbes. Les chercheurs allemands prévoient l'étendre à d'autres catégories grammaticales. En gros *Valenz* signifie la puissance d'un verbe à avoir des compléments. Connaissant la *Valenz* d'un verbe, on sait les possibilités de composer des phrases différentes avec ce verbe.

Notre conception de valence ne se limite pas aux seuls verbes, mais concerne toutes les catégories grammaticales. D'après Laforge (112, p. 186) on attribuerait la création de la notion de valence lexicale - en anglais *coverage* - à Ogden, en réaction consciente et arrêtée contre l'application "exagérée" du critère de fréquence en sélection. La valence est l'aptitude d'un mot à se substituer à un autre. C'est donc aussi la puissance de remplacement du mot; ce remplacement peut se faire par inclusion, par exclusion, par combinaison et par définition. Savard donne une explication détaillée de cette notion de valence dans son volume paru chez Didier en 1970, *La valence lexicale.*

Ce critère sert à mesurer l'utilité d'un mot; elle a surtout trait à l'économie lexicale. En effet un mot qui remplace plusieurs autres nous épargne l'apprentissage des longues listes. Des travaux d'analyse et de dépouillement ont été menés par Mackey et Savard (120) afin d'établir une liste comportant un indice numérique de valence pour chaque mot utile. Les résultats ont été obtenus en quantifiant chacun des quatre constituants du concept de valence et ont été publiés en 1967.

1.4.2.6 La familiarité

Les listes basées sur l'analyse statistique du discours semblent ne pas faire ressortir la façon dont les mots concrets sont emmagasinés dans la mémoire. Fraisse (51) dit notamment:

> "entre le système de la langue, socialement codifié et le monde des objets manipulés ou des situations vécues s'établit un tissu de relations de nature fonctionnelle. L'hypothèse se présente alors d'une donnée psychologique intermédiaire, à la fois reflet et incarnation de la fréquence, qui, en quelque sorte, ferait le pont

> entre le plan linguistique et le plan psychologique. Cette donnée nous a paru proche de ce que l'on appelle familiarité des mots. Cette familiarité n'est pas, comme la fréquence d'usage, un indice statistique de la langue, mais un trait qui caractérise pour chaque sujet, le langage qui est le sien."

Une étude de Richards (201) mesure l'importance générale des noms concrets de la langue anglaise, le même auteur vient d'écrire un travail sur la notion de familiarité comme critère de sélection du vocabulaire. L'auteur y fait d'abord une analyse critique de la fréquence, de la répartition, de la valence et de la disponibilité pour déterminer à quel degré ces critères résolvent les problèmes de la sélection du vocabulaire. L'auteur propose ensuite la notion de familiarité des mots et suggère que l'on tienne compte de l'expérience personnelle des individus dans le choix des éléments utiles:

> *"the choice of items for language teaching should not be limited to the use of the objective facts of language statistics revealed in frequency and range, or to the use of code centered analysis such as availability and coverage. Vocabulary selection should also take account of the facts of our experience, as they are reflected in subjective impressions of word familiarity"* (202, p. 132).

La même étude donne une description et les résultats de quelques travaux préliminaires pour mesurer la familiarité des mots. L'auteur présente ainsi une liste de mots rangés par degré de familiarité; cette liste pourrait servir au choix des mots utiles pour l'enseignement d'une langue seconde.

1.4.2.7 Facilité

Nous venons de rappeler dans les six points précédents quelques critères quantifiables qui donnent aux lexicométristes une indication sur l'utilité des mots. Un septième critère est celui de la facilité. C'est une notion que Laforge lie à l'économie d'effort. Elle a été appliquée dans le choix des éléments. Dubois, Drottens et Masserenti ont fait des études pour établir une échelle selon le degré de difficultés orthographiques. Pour l'instant on est arrivé à distinguer cinq facteurs ou (sous-) critères de facilité, ce sont: la similitude, la brièveté, la régularité, la clarté et la charge des éléments. Ces facteurs sont complémentaires aux critères d'utilité. La notion de facilité est traitée chez Laforge (112, chap. 8).

1.4.2.8 L'empirisme

Par empirisme en lexicométrie nous entendons ce critère de sélection le plus ancien sans doute qui consiste à choisir les éléments en se basant sur l'expérience mais aussi sur le hasard et le bon sens. Toute méthode qui n'utilise pas des critères objectifs est bien tenued'en trouver d'autres. La plupart choisissent alors l'empirisme. Quelques listes d'éléments ont été dressées au XXème siècle selon ce critère. Mais la tendance générale est de réduire l'utilisation de ce critère au profit des critères objectifs. Les défenseurs de ce mode de sélection soutiennent que l'empirisme permet de garder l'aspect personnifiant de la langue alors que les statistiques dépersonnifient la langue en quelque sorte.

1.4.2.9 Combinaison de deux ou plusieurs critères

Si la tendance initiale des chercheurs a été l'utilisation de la seule fréquence comme critère de sélection, on remarque depuis quelques décennies que la plupart des travaux optent pour une autre solution: ajouter un ou plusieurs critères à celui de fréquence.

Buchanan (13) désigne 5 324 mots d'indice minimum de fréquence 10 comme représentant l'essentiel de l'espagnol. Pour établir l'ordre de mérite des 2 000 premiers mots, il fait trois tests différents: un test de fréquence seule, un text de répartition seule, et un test de mérite où il combine les deux critères.

Garciá Hoz (55) tient aussi compte de la fréquence et de la répartition pour déterminer son vocabulaire fondamental espagnol.

Le Français Fondamental tient compte de la fréquence, de la répartition des mots et de la disponibilité. La liste a été augmentée de quelques mots choisis par la méthode empirique.

Savard et Richards (211) ont proposé l'utilisation de quatre critères objectifs. Partant des listes existantes, les auteurs utilisent la fréquence, la répartition, la disponibilité et la valence pour évaluer l'utilité d'environ 3 300 unités lexicales. Leur liste exclut les mots-outils, les auteurs prenant pour acquis qu'il s'agit là de mots utiles.

Le présent paragraphe vient de nous montrer une variété de critères de sélection, variété qui peut aussi entraîner une différence entre les compilations. Mais ce ne sont pas les seuls critères de sélection qui font cette différence entre les travaux, il y a aussi d'autres facteurs qui entrent en ligne de compte.

1.4.3 Corpus

L'importance du corpus est le troisième point que nous cherchons à observer dans l'analyse des travaux allemands. Dans la mesure où les données nous le permettent nous aimerions présenter le nombre total d'éléments qui ont été comptés; puis la quantité choisie, c'est-à-dire le nombre d'éléments différents trouvés après le dépouillement. Ces statistiques nous donneraient éventuellement quelques indications quant à la quantité d'éléments à prévoir pour préparer un vocabulaire limité d'avance. En d'autres mots si on voulait avoir un vocabulaire utile de 3 000 mots allemands, quelle quantité de matériel brut faudrait-il prévoir?

1.4.4 Méthode

Pour découvrir la méthode d'un auteur nous regarderons d'abord son échantillonnage. Il s'agira de voir le genre de textes ou de documents qui forment la base de son dépouillement. C'est ici que nous aborderons aussi la question de l'âge des textes, leurs catégories et leur grandeur. En second lieu nous examinerons les techniques de l'auteur. Son travail a-t-il été fait à la main ou à la machine?

1.4.5 Problèmes spécifiques

Au cours de sa recherche un lexicométriste peut se heurter à certains problèmes, nous réservons une rubrique à ces derniers. Dans cette section nous essayerons avant tout de dégager la définition que le lexicométriste donne à son unité de décompte. Certains classent les mots d'après leurs formes, d'autres d'après leur sens ou leur fonction. Nous nous demanderons aussi si l'auteur fait une différence entre adverbe, préposition et interjection; que fait-il des particules séparables? Les compte-il séparément ou les lie-t-il aux mots dont elles font partie?

1.4.6 Accessibilité

Il s'agit, sous ce point, d'étudier comment un auteur présente les résultats de ses travaux. Donne-t-il des listes alphabétiques dans lesquelles les mots sont suivis de leur fréquence, ou bien des listes de mots par rang de fréquence décroissante.

1.4.7 Innovation majeure

Au cours de l'analyse nous essayerons de relever les nouvelles notions et/ou les améliorations qu'un auteur aura apportées. Il s'agit d'innovation comme Keniston qui introduit la notion de ré-

partition, Michéa qui ouvre la porte à l'étude de la disponibilité; puis Savard et Richards qui proposent et quantifient l'utilité des unités lexicales suivant les quatre critères de base, fréquence, répartition, disponibilité et valence.

1.4.8 Utilisations ultérieures

Il va sans dire que plusieurs travaux des années cinquante et soixante n'ont pas encore connu d'utilisation. N'étant pas prophètes nous ne pourrons prédire ce que l'avenir leur réserve. Mais quant aux travaux de Kaeding, à ceux apparus au début du siècle ou même dans les années cinquante et qui ont été utilisés par d'autres chercheurs, nous essayerons de voir les différentes applications et/ou modifications qu'ils ont connues.

CHAPITRE II

ANALYSE DES TRAVAUX

2.1 Kaeding, F.W.

2.2 Les travaux basés sur Kaeding

2.2.1	Morgan, B.Q.	1928
2.2.2	Engel, E.F.	1931
2.2.3	Meier, Hermann	1931
2.2.4	Purin, C.M.	1931
2.2.5	Schinnerer, O. et Wendt, H.G.	1933
2.2.6	Vail, C.D.	1933
2.2.7	Bakonyi, H.	1934
2.2.8	Hohlfeldt, A.R.	1934
2.2.9	Wadepuhl, W. et Morgan, B.Q.	1934
2.2.10	Purin, C.M.	1937
2.2.11	Meier, Helmut	1952
2.2.12	Oehler, H.	1966

2.3 Autres travaux

2.3.1	Wadepuhl, W.	1928
2.3.2	Hauch, E.F.	1929
2.3.3	Simmons, L.V.T.	1929
2.3.4	New York State Syllabus	1930
2.3.5	Griesbach, H. et Schulz, D.	1958
2.3.6	Michéa, R.	1959
2.3.7	Uhlig, W. Chatelanat, C. et Lang, J.B.	1963
2.3.8	Wängler, H.H.	1963
2.3.9	Loriche, R.	1965
2.3.10	Melzer, F.	1965
2.3.11	Slaby, H.	1965

2.3.12 Wolf, F. 1969
2.3.13 Schmidt, P. 1970
2.3.14 Kopps, H. et Münster, T. 1971
2.3.15 Steger, H. et Keil, M. 1971

2.4 Pfeffer, J.A.

2.5 Tableau chronologique de la lexicométrie allemande

Si l'établissement de la documentation bibliographique était difficile dans cette étude, l'accès aux ouvrages l'a été encore plus. En dépit de maints efforts il ne nous a pas été possible de mettre la main sur bon nombre de volumes. Mais les bibliographies annotées nous ont permis d'avoir une certaine idée des travaux que nous n'avons pas pu consulter directement. Les vingt-neuf analyses qui vont suivre sont loin de la totalité des recherches faites en lexicométrie allemande, mais nous croyons qu'elles représentent assez bien la situation de cette branche de la lexicologie pour la langue allemande.

2.1 Kaeding, F.W.

Buts

Le *dictionnaire de fréquence* de Kaeding (102) est souvent considéré comme l'ancêtre des statistiques de vocabulaire, car il représente le premier travail méthodique d'envergure. L'ouvrage est destiné à l'établissement d'une sténographie rationnelle. L'auteur cherche un nouveau système pour améliorer la notation sténographique. Il s'agit donc avant tout d'un travail qui vise un résultat pratique et utilitaire.

Critère de sélection

Kaeding choisit la fréquence des unités de décompte comme critère de sélection.

Corpus

L'auteur avait préparé 110 sous-ensembles (ou échantillons) de textes de 100 000 mots chacun; mais le chiffre final est de 10 910 777 occurrences prises dans 14 sources différentes. De tous ces mots il n'en retient que ceux ayant un indice de fréquence de 4 ou plus, ce qui lui donne un total de 79 716 mots différents. Les 14 sources dépouillés peuvent être regroupées et réduites comme suit:

journaux et revues	39.4% du total des mots courants
textes administratifs	3.7% du total des mots courants
ouvrages de théologie	3.7% du total des mots courants
ouvrages de médecine	0.9% du total des mots courants
textes commerciaux	2.8% du total des mots courants
textes historiques	4.6% du total des mots courants
classiques allemands et étrangers	18.3% du total des mots courants

correspondance commerciale, militaire et personnelle	9.2% du total des mots courants
débats parlementaires	8.2% du total des mots courants
livres scientifiques	3.7% du total des mots courants
documents militaires	5.5% du total des mots courants

Ce corpus de Kaeding peut être qualifié de très représentatif. En effet, l'auteur essaye d'y inclure les genres de textes écrits que l'on pouvait trouver de son temps.

Méthode

Le dépouillement se fait à la main. Le travail nécessite le concours de 5 000 sténographes et de 800 collaborateurs. Les mots étrangers sont comptés, mais pas ceux des longues citations, car on voulait traiter les phrases allemandes et non étrangères. Les noms géographiques, les noms propres, les chiffres et les titres de chapitre sont exclus. Seul le texte suivi est pris en considération. L'orthographe adoptée était celle proposée par le Duden d'alors, une grammaire orthographique allemande.

Problèmes

C'est la forme du mot qui représente l'unité de décompte chez Kaeding. Un singulier et son pluriel seront donc comptés comme deux mots différents. Ceci est normal pour un travail destiné à la sténographie. Kaeding considère le mot comme une combinaison particulière de lettres, de syllabes ou de sons; il ne tient pas compte du sens ni de la construction grammaticale. Il définit ce qui importe pour le sténographe en ces termes:

> *"Den Stenographen liegt daran, festzustellen,*
> *1. Die Häufigkeit der einzelnen Wörter*
> *2. Die Häufigkeit der einzelnen Silben*
> *3. Die Häufigkeit der einzelnen Laute*
> *4. Die Häufigkeit der einzelnen Buchstaben*
> *und zwar*
> *a) der einzelnen Buchstaben*
> *b) der Zusammensetzungen*
> *c) der Verschmelzungen."* (102, p. 10, version 1898)

Ces quatre points ont été visés au cours de la recherche.

Le dictionnaire de Kaeding présente un certain nombre d'ambiguités: par exemple les chiffres de fréquence entre parenthèses indiquent qu'un mot est apparu avec une première lettre majuscule. Sans que l'auteur du travail ne le précise, on suppose seulement qu'il s'agit des substantifs; toutefois ceci ne peut être justifié car après un point le mot allemand commence aussi par une majuscule.

Accessibilité

Le travail comprend des listes de mots, de suffixes, et de lettres. Les résultats des dépouillements sont présentés sous forme de listes alphabétiques dans lesquelles un mot est suivi d'une série de chiffres. On aura par exemple

arme (740) 1362 / 1423

ce qui signifie que le mot *arme* apparaît comme forme isolée et comme partie d'un mot composé en tout 1423 fois. Il apparaît comme forme isolée 1362 fois et il revient avec une lettre majuscule *Arme* 740 fois.

La toute première liste donne les mots dont l'indice de fréquence est 5 000 ou plus.

La seconde range les mêmes mots par ordre de fréquence décroissante avec l'article *die* en tête 358 054 et *wenig* qui clôt la liste avec un indice de 5 001. Ce sont les mots grammaticaux qui occupent les premiers rangs. Le premier substantif vient au 77e rang avec une fréquence de 21 596 *Zeit*. Le second substantif occupe le 103e rang avec une fréquence de 14 416 *Ordnung*. Dans les 100 premiers mots il n'y a qu'un seul substantif et 4 infinitifs *nehmen*, *werden*, *sein* et *haben*. La plupart des mots n'ont qu'une syllabe, le premier à deux syllabes est la préposition *über* qui vient au 26e rang. Des 320 mots de la première liste il y en a 171 à une syllabe, 141 à deux syllabes et 8 à trois syllabes. Il y a alors lieu de croire au principe du moindre effort.

La troisième liste est alphabétique et donne les mots dont l'indice de fréquence varie de 4 à 4 999.

Vient ensuite une liste alphabétique des radicaux d'indice de fréquence égal ou supérieur à 5 000. Chaque forme y est suivie de deux chiffres dont le premier désigne la fréquence totale et le second la fréquence moyenne sur 100 000 syllabes. La syllabe étant une unité importante pour le sténographe, Kaeding étudie aussi la fréquence des radicaux par rapport aux syllabes. Son corpus compte 20 000 000 de syllabes, mais il ramène les résultats en étudiant la fréquence moyenne dans 100 000 syllabes. C'est ainsi que la forme *Frau* apparaît comme radical

7 735 fois sur les 20 000 000 de syllabes du corpus et en moyenne 38,67 sur 100 000 syllabes.

La cinquième liste donne les mêmes radicaux rangés par ordre de fréquence décroissante.

C'est dans la sixième liste qu'apparaissent les radicaux dont l'indice de fréquence descend jusqu'à 8.

L'étude de Kaeding a également porté sur les particules. Aussi en donne-t-il les résultats dans la septième liste où il présente les particules et leurs combinaisons. L'auteur fait une distinction entre particules séparables et particules inséparables. Seules les premières ont été isolées comme formes et comptées comme telle. La liste A donne le nombre de fois que la particule a été rencontrée seule ou en combinaison. La liste B détermine les liaisons de chaque particule. La liste C range les particules par ordre de fréquence. Les listes D et E donnent respectivement les particules séparables et inséparables. La liste F fournit les terminaisons des particules.

Une huitième section est réservée aux suffixes:

A - Keading donne la liste des suffixes qui apparaissent plus de 1 000 fois seuls ou en liaison:

ex. *ung* 165 856 / 236 079 1 180,39

Le premier chiffre est l'indice de fréquence d'apparition seule, le second, l'indice totale de fréquence d'apparition seule et en combinaison et le dernier chiffre, la fréquence d'apparition sur 100 000 syllabes.

B - Cette liste indique tous les suffixes par ordre alphabétique. Les chiffres qui accompagnent les mots représentent le nombre de combinaisons qu'un suffixe peut avoir.

C - Dans cette dernière liste Kaeding relève les terminaisons qui ne se rencontrent jamais seules.

Ces huit listes représentent la première partie du *dictionnaire de fréquence*. La deuxième partie est présentée aussi sous forme de tableaux dont le premier traite du décompte des graphèmes (*Buchstabe*).

Les 20 000 000 de syllabes allemandes du corpus ont 47 479 805 de formes et 60 558 018 de graphèmes. Les résultats de Kaeding montrent qu'une syllabe allemande a en moyenne 3,03 graphèmes.

Le second tableau donne la fréquence de toutes les formes. Connaissant la fréquence des différents graphèmes, Kaeding s'est intéressé à étudier l'entourage de chacun d'eux et il en donne les résultats au troisième tableau. Par exemple la graphie *æig* (102, p. 542) s'est trouvée.

en début de syllabe	1 505 fois
après b	299 fois
après ch	76 fois
après d	1 fois
après dsch	1 fois

Le quatrième tableau donne les formes de consonnes puis les voyelles avec lesquelles on les rencontre. Pour le groupe des consonnes 'dsch' par exemple, les recherches de Kaeding ont prouvé qu'il n'était suivi de la voyelle 'a' qu'une seule fois.

Le tableau cinq apporte la statistique des sons. Il présente les sons avec leur fréquence ainsi que la forme dans laquelle ces sons apparaissent.

Une dernière liste donne les syllabes dont l'indice de fréquence se situe entre 1 et 7.

Le travail de Kaeding inclut également l'analyse de la ponctuation. Il voulait relever le genre de ponctuation qu'on rencontre en allemand.

L'accessibilité de ce dictionnaire est sans doute bonne pour le sténographe mais pas tellement pour un professeur de langue. En effet il n'y a pas moyen de déterminer avec précision combien de fois un mot comme *Band* apparaît avec le sens 'ruban', 'lien moral', ou 'volume'. Un mot et son pluriel ne seront pas près l'un de l'autre puisque les mots sont rangés par ordre alphabétique et les voyelles avec *Umlaut* sont considérés être suivies d'un e muet - ü devient alors ue, ä - ae et ö - oe... -. On trouvera donc à la page 262 le mot *Mütter* (pluriel de 'mère') 85/102 et à la page 263 le mot *Mutter* (= 'la mère') 3 744/4 693. Les formes d'un même verbe sont également éparpillées dans la liste.

En dépit de ces difficultés ainsi que de l'âge des textes l'oeuvre de Kaeding représente encore un document utilisable à d'autres fins (voir "Utilisations ultérieures" plus loin).

Kaeding établit la technique de dépouillement sur laquelle se baseront la plupart des recherches ultérieures. Il donne un bon exemple d'échantillonnage. Il puise à diverses sources. Celles-ci représentent la quasi totalité du matériel écrit que l'on pouvait trouver de son temps.

Utilisations ultérieures

Le dictionnaire de fréquence a été utilisé à maintes reprises pour la préparation d'autres listes de mots. Les oeuvres qui suivent en font un témoignage. Parmi celles-ci se classent aussi la *Semantic Frequency List* de Eaton (35). Eaton utilise les listes existantes dont celle de Kaeding pour dresser un vocabulaire d'environ 6 500 concepts rangés par sections de 1 000 éléments selon la fréquence des mots dans les quatre langues suivantes: l'anglais, le français, l'allemand et l'espagnol.

Henmon et al. (90) préparent aussi un test de langue en 1926. La partie vocabulaire de ce test utilise les travaux de Kaeding. En effet les membres du conseil choisissent systématiquement 100 mots du *dictionnaire de fréquence*.

2.2 LES TRAVAUX BASES SUR KAEDING

Sous cette rubrique nous comptons englober les ouvrages dont le vocabulaire est tiré directement ou indirectement de la recherche de Kaeding. Certains auteurs ont utilisé directement le *Dictionnaire de fréquence* en choisissant quelques éléments lexicaux pour dresser leur vocabulaire de base. D'autres par contre se sont servis non pas du *dictionnaire de fréquence*, mais des listes basées sur celui-ci. Dans un cas comme dans l'autre il s'est trouvé quelques chercheurs qui ont ajouté au vocabulaire ainsi récupéré, d'autres éléments lexicaux choisis ailleurs ou par des recherches personnelles.

2.2.1 Morgan, B.Q. *German Frequency Word Book* 1928

Buts

Morgan vise l'enseignement de l'allemand écrit et l'acquisition d'un vocabulaire pour lire des textes. Il dit dans sa préface:

> *"This work provides a list of words of the greatest importance for learners and teachers, for the making of grammars, and for the selection of reading texts... The lists have an importance for teaching and for the acquisition of a vocabulary for reading German which is unequalled."* (157, V et VI)

Critère de sélection

La liste de Morgan étant basée sur celle de Kaeding, la fréquence reste le seul critère fondamental ayant servi au choix des éléments. Morgan apporte cependant une légère modification en ce sens qu'il utilise la fréquence groupée comme on le verra au paragraphe suivant.

Méthode

Il additionne les chiffres de fréquence des différentes formes puis il les regroupe sous des entrées principales ou mots clefs. Un infinitif se retrouve donc avec toutes ses formes relevées, un singulier et son pluriel ainsi que leurs formes déclinées sont regroupées. Morgan obtient ainsi une première liste de base dont les mots ont chacun un indice de fréquence d'au moins 200. Le second groupe est une simple liste alphabétique d'environ 6 000 mots. Cette liste contient les mots de base ainsi que leurs dérivés communs dont l'indice de fréquence est égal ou supérieur à 100 dans le *dictionnaire de fréquence*.

Un trait caractéristique de la liste de Morgan est la notion de radical *basic or stem word* et de mot dérivé *related or derived word*. Cette distinction a déterminé la forme de sa compilation.

Problèmes

Le problème majeure a été de rendre le travail de Kaeding utilisable dans l'enseignement des langues. Morgan arrive à le résoudre en introduisant ces notions de mots de base et mots dérivés. Ces notions conduisent à l'affirmation que tous les mots qui associent de façon reconnaissable des formes congénères contribuent à fixer le sens du mot de base dans l'esprit de l'étudiant.

Accessibilité

La liste A range les 2 402 mots par ordre de fréquence décroissante. Morgan n'indique pas la fréquence du mot mais plutôt l'intervalle de fréquence dans lequel ce mot se trouve. Il fait seulement précédé chaque élément d'un numéro d'ordre. Aussi aura-t-on le mot 18 *nehmen* dont le numéro d'ordre est 18 dans l'intervalle 99 999 - 50 000.

Dans la liste B les mots dérivés sont rangés par ordre alphabétique sous le mot de base correspondant. Sous le mot *bilden* de fréquence 18 on trouve: *ausbilden, Ausbildung, Bildung, nachbilden, Nachbildung*, et *Vorbildung* (157, p. 31).

Innovations

Un des grands mérites de Morgan est sans doute d'avoir transformé le travail de Kaeding en un document utilisable en didactique des langues. Mais il nous présente aussi une autre définition du mot. Ce n'est plus une simple forme comme le voient les sténographes; le mot est lié à la notion du concept sémantique: C'est ce qui explique le regroupement des fréquences des différentes formes d'une même racine sous un mot clef.

C'est de Morgan que nous apprenons les notions de *basic or stem words* et *related or derived words*.

La liste de Morgan est essentiellement adaptée aux adultes. En effet les mots les plus fréquents se rapportent surtout au gouvernement, aux finances, à l'industrie, à la jurisprudence et à l'impôt alors que les termes de la vie courante sont plutôt rares.

L'âge du vocabulaire se reconnaît par le fait que des mots comme *König* (roi), *Reich* (empire) s'y trouvent avec un indice de fréquence assez important, alors que d'autres comme *automobile* manquent.

2.2.2 Engel, E.F. *The Usage of a Standardized Vocabulary in Beginning German* 1931

Selon Coleman, Engel critique quelques listes dont celle de Wadepuhl (1928) et celle de Morgan (1928) auxquelles il reproche d'avoir omis les mots de la vie quotidienne. L'auteur rapporte que dans son école, 1 450 mots sont maîtrisés au bout du 1er semestre. Il ajoute 605 autres mots fondamentaux qu'il choisit dans les listes existantes mais aussi de façon subjective.

2.2.3 Meier, Hermann *The 1 000 Most Frequent German Words* 1931

But

La liste est préparée pour l'enseignement de la langue allemande.

Méthode

L'auteur tient compte de la liste de Morgan, aussi signale-t-il les 500 premiers mots de cette liste se trouvant dans la sienne.

Présentation

Les congénères apparaissent en lettres grasses avec indications de la différence entre les congénères anglais. Les 1 000 mots les plus fréquents sont suivis de leur traduction en anglais. Quelques pages à la fin du livre contiennent des renseignements sur la façon d'apprendre, sur la prononciation et la grammaire. Pour cette dernière, il y a des indications sur les articles, les terminaisons d'adjectifs, les pronoms et les formes verbales.

2.2.4 Purin, C.M. *A Standard German Vocabulary* 1931

But

L'auteur vise l'enseignement de l'allemand. Coleman dit en l'occurence:

> *"The list is an interesting attempt to exploit the frequency counts of the past decade for teaching purposes."* (22, p. 130)

Méthode

Cette liste est dérivée des recherches basées sur le critère de la fréquence. Mais la sélection de Purin n'est pas tout à fait objective. En effet voici ce qu'en dit Coleman:

> *"very frequently the inclusion of a word was determined simply by its usefulness in narration of incidents of common occurrence."* (197, p. 130)

En préparant sa liste, Purin se laisse guider par l'apparition ou non apparition d'un mot donné dans une douzaine de textes les plus largement utilisés. Il compare ensuite sa liste avec les 1 000 mots de la liste de Wadepuhl (1928). Les tournures idiomatiques sont tirées du dictionnaire de Hauch (2.3.2). Purin traduit les mots et les tournures idiomatiques, et ajoute d'autres détails comme la flexion des noms et des verbes.

2.2.5 Schinnerer, Q. et Wendt, H.G. *A Suggested List of 1 000 Active German Words* 1933

Buts

Les auteurs visent les mots de la langue courante et dressent une liste d'un vocabulaire actif pour l'apprentissage de l'allemand.

Corpus

Ils se fixent un maximum de 1 000 mots pour leur liste finale.

Méthode

Les auteurs incluent dans leur liste, des articles, des pronoms, adjectifs possessifs, numéraux, les jours de la semaine, les mois et d'autres mots communs tel que *Kreide, Tafel, Fahrkarte, Strassenbahn, Konzert.*

Pour éviter d'être entièrement subjectifs, ils consultent aussi quelques listes existantes, entre autres Morgan-Kaeding (2.2.1), Meier Hermann (2.2.3), Purin (2.2.4), Engel (2.2.2).

Présentation

Par des tableaux les auteurs présentent le nombre de mots de chaque liste consultée, le nombre de mots de leur liste tirés de chacune des listes précédentes puis le nombre de mots non trouvés dans ces listes mais qui se retrouvent dans la leur.

2.2.6 Vail, C.D. *Basic Word and Idiom Liste* 1933

a) *Basic Word List*

But

L'auteur vise l'enseignement de l'allemand et il cherche à établir des vocabulaires de base pour les trois premières années d'étude.

Méthode

Il dépouille sept grammaires et six *readers* les plus populaires. Mais il tient également compte des six listes déjà existantes. Du total des mots ainsi obtenus, l'auteur dérive trois listes contenant respectivement 760 mots pour la première année, 754 pour la deuxième et 598 pour la troisième.

Par vocabulaire de base l'auteur entend donner un minimum de mots que chaque élève devrait connaître et être prêt à utiliser. Cela ne veut pas dire que l'élève, le professeur ou l'auteur des manuels doivent se limiter strictement à ces quelques mots.

b) *Basic German Idiom List*

Préparé par C. Schultz ce travail est identique au précédent. D'un corpus de cinq livres de lecture, l'auteur sélectionne environ 400 tournures idiomatiques d'après le critère de la fréquence. Il les compare aussi avec les tournures idiomatiques de la liste de Hauch (85).

2.2.7 Bakonyi, H. *Die gebräuchlichsten Wörter der deutschen Sprache* 1934

But

L'auteur cherche à promouvoir l'enseignement de l'allemand comme langue étrangère et son vocabulaire se destine aux débutants. Il recherche la langue parlée quotidienne et la langue littéraire simple.

Critère

Bakonyi utilise un critère différent de celui de ses prédécesseurs. Il veut chercher les *gebräuchlichste Wörter* - que nous rendons par mots usuels. Pour lui un mot usuel est un mot qui est fréquent et connu (*gebräuchlich = häufig + bekannt*). Un mot fréquent est généralement connu, mais un mot connu n'est pas forcément fréquent. Pour comprendre le terme usuel chez Bakonyi il nous semble qu'on devrait pouvoir dire aussi qu'un mot usuel est un mot connu et fréquent. Bakonyi fait une distinction entre mots usuels et mots fréquents et semble avoir aperçu, sans toutefois le mettre en relief, la notion de disponibilité (1.4.2.4). Pour lui l'étude de la fréquence indique le nombre de fois qu'apparaît un mot. Mais un dictionnaire des mots les plus connus cherche le nombre de mots différents dans un corpus, sans s'intéresser au nombre de fois qu'ils se répètent. En fréquence on considère la répétition alors qu'ici c'est le nombre de *Wortvorräte* (7, p. 11) (réserve des mots) qui importe. Bakonyi aperçoit aussi la notion de répartition (1.4.2.2). En accordant une certaine valeur (numérique) à ces mots, il dit en l'occurrence:

> *"In je mehr Quellenwerken ein Wort vorkommt, desto grösser ist im allgemeinen auch dessen Wertstufe."*

(Plus un mot se retrouve dans plusieurs sources, plus il a de la valeur.)

Corpus

Bakonyi se choisit 19 sources composées de plusieurs listes préétablies et totalisant environ 30 000 mots. Le dépouillement donne 7 530 mots différents (*Einzelwörter*), d'où l'auteur tire 5 874 mots usuels, c'est à dire les plus connus et les plus fréquents.

Méthode

Pour trouver les mots les plus connus, Bakonyi choisit 19 centres d'intérêt répartis en trois groupes. Une première section se compose de textes de langue allemande pour les débutants. Ces textes sont les plus nombreux. Une deuxième partie, moins grande en quantité, mais plus importante pour l'auteur, se compose d'une part de listes de mots pour enfants de 2 à 6 ans et d'autre part des listes de mots pour sourds-muets.

Le travail de Bakonyi se fonde principalement sur l'analyse de la langue des enfants. L'auteur estime que le parler le plus simple qui utilise les mots les plus connus est celui des enfants. Il constate que les sourds-muets sont des personnes qui apprennent peu de mots et ce peu de mots ils l'apprennent difficilement; les pédagogues devaient donc leur présenter des mots plus faciles et utiles. En choisissant ses mots parmi ceux des sourds-muets, Bakonyi croit donc trouver les mots les plus simples et les plus connus de l'homme normal. Les mots fréquents sont pris de la liste de Kaeding.. Plus le mot est usuel pour un jeune ou pour un moins instruit, plus ce mot est connu. Tel est le principe qui guide Bakonyi dans l'évaluation des mots de son corpus. Il attribue à chacune de ses 19 sources un coefficient numérique; les plus forts étant réservés aux mots des enfants et des sourds-muets. S'il avait donné à chaque source le coefficient 1 par exemple, le plus haut indice d'un mot serait 19, mais les coefficients étant différents l'indice de mesure le plus élevé est 36. Bakonyi se base sur cet indice pour présenter les résultats de son travail.

Accessibilité

L'auteur donne quatre tableaux. Le premier présente 1 470 mots dont l'indice se situe entre 36 et 11. La deuxième partie du même tableau donne les mêmes mots par ordre alphabétique puis avec indication du degré de fréquence dans la liste de Kaeding. Les mots de Kaeding ont été répartis en trois groupes de coefficient

3 pour les mots de fréquence 358 054 - 2 000
2 pour les mots de fréquence 2 000 - 500
1 pour les mots de fréquence 500 - 200.

On lira donc (7, p. 37) 3 *auf*: 36, ce qui signifie que la préposition *auf* a une fréquence comprise entre 358 054 et 2 000 dans la liste de Kaeding, son indice dans la liste de Bakonyi est 36. Il s'agit donc d'un mot très usuel.

Le deuxième tableau donne les 1359 mots d'indice 10 - 6. Ils sont classées par ordre alphabétique avec indication de l'indice d'évaluation.

Le troisième tableau donne les 1 585 mots d'indice 5 - 3 et enfin, le quatrième présente 1 460 mots dont l'indice est 2.

Innovation

Bakonyi se base donc sur son critère de mot usuel pour présenter les résultats de sa recherche. Ce critère constitue un aspect nouveau en lexicométrie. Le terme usuel chez Bakonyi pourrait peut-être s'appliquer à un mot à la fois fréquent et disponible si nous utilisions le langage lexicométrique moderne.

Un autre aspect nouveau dans le travail de Bakonyi est la recherche des mots usuels dans le langage des enfants. Cette méthode repose sur un raisonnement très logique: tout homme ayant été enfant a d'abord connu un vocabulaire restreint qui, en dépit des différences provenant des milieux, des niveaux de vie, des sociétés où l'enfant grandit, est quand même commun à tous et par ce fait même est connu de tous les adultes. Un étranger qui apprendrait alors ce vocabulaire pourrait bien se faire comprendre dans sa langue seconde.

2.2.8 Hohlfeld, A.R. *Minimum Standard Vocabulary German* 1934

Buts

La liste vise l'enseignement et l'apprentissage de l'allemand.

Critère de sélection et méthode

Il s'agit d'une liste combinée ayant comme base la liste de Morgan (2.2.1) d'où Hohlfeld élimine certains termes techniques pour les remplacer par quelques mots disponibles *utility words* choisis dans les listes de Vail (2.2.6) et de Purin (197). Le comité qui fait ce travail obtient alors une liste de 2 120 mots que les étudiants devaient maîtriser à la fin d'un 4ème semestre d'étude. Les noms des jours, des mois, des nombres et les pronoms sont exclus de la liste mais le professeur est censé les enseigner.

Accessibilité

Les mots qui devaient être appris en deux semestres sont en caractère gras. Les mots dérivés viennent derrière le mot clef comme dans l'exemple suivant:

sammeln: gesamt, Gesamtheit, Sammlung, samt, sämtlich, versammeln, Versammlung, zusammen.

Innovation

Le fait de distinguer entre les mots devant être connus en un certain temps (deux semestres ou quatre semestres) dénote une orientation vers la gradation des éléments à enseigner.

Le regroupement des mots autour d'un mot clef se retrouve aussi dans le travail de Hauch (2.3.2) et montre cette croyance implicite à la notion de champ sémantique.

2.2.9 Wadepuhl, W. et Morgan, B.Q. *Minimum Standard German Vocabulary* 1934

C'est sur la demande du *American Association of Teachers of German (AATG)* qu'un comité est chargé d'étudier la question du vocabulaire fondamental allemand et d'en faire rapport. Les membres de ce comité dont Morgan (directeur) et Wadepuhl faisaient partie préparent ce rapport ainsi qu'une liste de mots qui sont acceptés par l'association. Le travail comporte deux parties, une simple liste de mots puis un dictionnaire des équivalents anglais de ces mots.

But

Le dictionnaire est destiné à l'enseignement de la langue et doit surtout aider deux groupes de personnes:

a) l'étudiant pour qu'il lise des textes allemands et

b) le professeur et l'auteur des livres de lecture qui ont besoin de connaître ce à quoi on doit s'attendre normalement d'un élève dans son bagage lexical.

Critère et méthode

La liste des mots est un sous-produit de la liste de Kaeding en ce sens qu'elle est faite à partir du *German Frequency Word Book* de Morgan (157), ce dernier ouvrage se basant sur le *dictionnaire de fréquence* (102).

L'arrangement des paragraphes du dictionnaire correspond au principe des mots de base et dérivés tel que décrit dans la préface du *German Frequency Word Book* et retenu comme principe de travail par le comité.

Corpus

Le nombre de mots de base s'élève à 2 150. Ce total constitue, pour le comité, un minimum à compléter après 4 ans d'étude au *High-School* ou après un cours de 2 ans au *College*. De ce total 1 018 mots constituent un minimum à maîtriser après 2 ans de *High-School* ou 1 an du *College course*. Les mots de ce dernier total sont précédés d'une étoile dans la liste.

Accessibilité

Les mots présentés dans chaque paragraphe du dictionnaire sont groupés en famille de mots. Tous sont rangés par ordre alphabétique si bien qu'on peut remarquer du coup si un mot particulier se trouve dans le dictionnaire ou non.

Un appendice donne une liste spéciale des verbes forts, en indiquant successivement la 3ème personne du singulier du présent de l'indicatif, du prétérit et du passé du subjonctif, puis l'impératif et le participe passé. Ces détails grammaticaux sont utiles aux étudiants d'allemand.

2.2.10 Purin, C.M. *A Standard German Vocabulary* 1937

But

Ce dictionnaire est destiné à l'enseignement de l'allemand. L'auteur dit à ce propos dans la préface:

> *"It (the wordbook) should.....amply meet the basic vocabulary needs of four-year courses in secondary schools and of elementary and intermediate courses in colleges covering the first four or five semesters."* (198, p. iii)

Critère et méthode

Il s'agit d'une liste combinée. Purin étudie préalablement les publications sur la fréquence des mots notamment celles de Wadepuhl-Morgan (2.2.9) et Vail (249); mais il se sert aussi des résultats de son travail de 1931 (2.2.4). Dans ce dictionnaire l'auteur exclut un certain nombre d'éléments: les noms propres, les noms des mois, les pronoms personnels et possessifs ainsi que les adjectifs possessifs, il en est de même de tout chiffre ou adjectif ordinal et cardinal, des mots étrangers ayant gardé leurs formes et conservé leur sens en allemand et en anglais (comme *Tennis, General, Geographie, Museum...*). Sont aussi exclus les formes déclinées des articles, les mots composés dont une par-

tie est facile à comprendre - comme *Haustür, Klassenzimmer*..... -, un nombre considérable de noms et d'adjectifs avec le préfixe de négation *un-* et dont le sens est très clair pour l'étudiant qui connaît le sens du mot affirmatif. Si le mot *Sinn* (sens) est connu, le mot *Unsinn* ne sera plus dans la liste.

Se fondant sur l'expérience des professeurs d'allemand, Purin inclut un certain nombre de mots utiles que les deux listes (Wadepuhl et Vail) ne contenaient pas.

De 1 000 premiers mots les plus fréquents de la liste Wadepuhl-Morgan 967 ont été retenus dans le dictionnaire de Purin.

Corpus

Ce dictionnaire contient 2 932 mots classés par ordre alphabétique, environ 2 000 mots dérivés et quelques 1 500 tournures idiomatiques.

Accessibilité

Les 500 mots les plus fréquents sont précédés de deux astérisques. Chaque mot a une traduction anglaise, sauf ceux pour lesquels un synonyme allemand ou une paraphrase peut éclaircir le sens. Comme le verbe *erwerben* peut être remplacé par son synonyme *verdienen* il n'aura pas de traduction. Pour aider l'étudiant à retenir un mot le congénère anglais est souvent mis entre parenthèses:

erscheinen (*ie, ie*) (cog. *shine*) *appear*... (198, p. 44)

Les dérivés les plus fréquents sont indiqués entre crochets:

der Diener, -s, - (dienen) servant (masc.)
[*die Dienerin - servant (fem.)*]
der Dienst - service, duty.] (198, p. 30)

La plupart des mots sont illustrés par des exemples. Ce qui permet aux étudiants d'apprendre les mots dans un contexte. Les mots sans exemples sont généralement les noms concrets. Pour un mot ayant plusieurs sens, seuls les sens les plus fréquents sont illustrés par des exemples.

Purin augmente sa première liste d'environ 1 000 nouveaux termes. Son travail est utile pour les étudiants à cause des exemples et de nombreux dérivés qu'il inclut. On n'apprendra plus de listes de mots par coeur, mais des mots dans des phrases.

2.2.11 Meier, Helmut *"Die tausend häufigsten Wortformen der deutschen Sprache"* 1952

Méthode

Se basant sur le compte des syllabes de Kaeding (2.1), Meier dresse une liste de 1 022 formes de mot les plus fréquentes. Il en indique aussi l'indice de fréquence.

2.2.12 Oehler, H. *Grundwortschatz Deutsch* 1966

But

L'auteur vise l'enseignement de l'allemand comme langue étrangère.

Critère

Il établit son allemand fondamental grâce aux recherches effectuées dans différents pays sur le plan de la fréquence des mots. Mais Oehler fait aussi intervenir la méthode empirique avec la participation des professeurs; il dit à ce propos:

> *"Die Häufigkeit konnte bei unserer Auswahl nicht einziges Kriterium sein. Neben dem Sprachstatistiker hat der Sprachpädagoge sein Wort gesprochen; aus seiner Lehrerfahrung konnte und musste er ergänzen und auch berichtigen."* (177, p. 231)

Ce que nous ignorons c'est le pourcentage de mots introduits par cette méthode empirique.

Corpus

Son vocabulaire de base contient plus de 2 000 mots et 3 000 tournures idiomatiques de la langue allemande avec leurs traductions en anglais et en français.

Méthode

L'ouvrage est censé refléter la langue moderne parlée et écrite. Il contient les résultats des recherches de Kaeding, Bakonyi, Schultze, Michéa, Wängler, Meier et Pfeffer. En dressant sa liste, Oehler laisse de côté les mots qui n'ont pas de correspondants dans les deux autres langues. Il exclut aussi les noms de personnes, de villes, des animaux, plantes, boissons, puis des maladies pour lesquelles il ne se trouve pas

de concepts équivalents en français et en anglais ou des termes qui ne peuvent être reconnus comme appartenant véritablement au vocabulaire de base allemand, tels que *doll, Sekt, Forelle, Veilchen, Knecht, Masern.*

Accessibilité

La liste se présente sous forme d'un dictionnaire normal. Plusieurs détails utiles y apparaissent: l'accent, les antonymes, les inflexions, les temps primitifs des verbes forts, l'indication sur la possibilité de séparation des particules.

Le dictionnaire est suivi de deux index dont un Anglais-Allemand et l'autre Français-Allemand.

Utilisations

Ce travail est utilisé par Mattutat (128) pour présenter les possibilités d'emploi du vocabulaire fondamental allemand. Il s'agit d'un ouvrage qui est destiné à l'autochtone qui veut approfondir les connaissances de sa langue maternelle, mais aussi à l'étranger non débutant qui veut améliorer sa compréhension de l'allemand et élargir ses possibilités d'expression.

Mattutat divise son travail en trois parties:

a) les mots grammaticaux;

b) la liste des mots par ordre alphabétique;

c) la syntaxe de l'allemand.

L'auteur donne quelques indications sur l'ordre des mots dans une phrase allemande.

2.3 AUTRES TRAVAUX

Il s'agit avant tout de travaux qui n'ont pas utilisé les résultats de Kaeding pour dresser leur liste. Il y en a très peu avant les années cinquante. Ces travaux n'abondent qu'après la deuxième moitié du XXe siècle. Ceci ne voulant pas du tout dire qu'on ait oublié Kaeding.

2.3.1 Wadepuhl, W. *German Idioms* 1928

L'auteur dresse deux listes de 100 tournures idiomatiques chacune. Il s'agit de tournures apparaissant assez fréquemment dans des textes utilisés par les étudiants de 1ière et 2ième année d'allemand. Ses deux listes sont destinées respectivement aux étudiants de première et deuxième année. Wadepuhl cite les livres d'où il a extrait ses tournures.

2.3.2 Hauch, E.F. *German Idiom List* 1929

But

Le livre est destiné à l'enseignement de l'allemand comme langue écrite. L'auteur du travail précise dans l'introduction:

> *"The purpose for which the German idiom count was undertaken was to provide a preferential list that teachers and writers of textbooks could use as a concrete guide in the selection of classroom and textbook material."*

plus loin, il ajoute (85, p. 3):

> *"The count... applies specifically to the language as written rather than to the spoken language."*

Critères

Pour choisir les idiomes, Hauch fait un décompte en se basant sur les critères de la fréquence et de la répartition. Il faut cependant remarquer que la fréquence l'emporte sur la répartition, car c'est elle qui représente le critère principal dans le regroupement des tournures idiomatiques. La fréquence varie de 225 à 3 - soit respectivement les mots *gar* 225/25 et *ob* 3/1. La répartition varie de 43 à 1; le mot dont l'indice de répartition est le plus élevé est *geben* (*es gibt*) 211/43.

Corpus

Le décompte est fait sur un corpus totalisant environ un million d'occurrences. Le résultat donne une liste de 959 tournures idiomatiques rangées d'après leur mot clef.

Méthode

Hauch choisit une population lexicale de 57 textes (ou échantillons). Il s'agit surtout de textes de prose mais aussi de théâtre du XVIIIe au XXe siècle. La répartition est approximativement la suivante:

textes du XVIIIe et début du XIXe siècle	20% du total
textes de 1820 — 1870	20%
_____ 1870 — 1900	30%
_____ du XXe siècle	30%

Les 57 textes sont de longueur inégale, mais varient dans l'ensemble entre 7 000 mots et 25 000. Pour les dépouillements, Hauch s'est fait aider par 46 membres du corps enseignant.

Problèmes

L'auteur définit tournure idiomatique de façon générale comme:

> *"A combination of words that upon analysis for vocabulary and syntactical characteristics conveys either no intelligible meaning at all or one differing from the conventionally accepted sense."* (85, p. 1)

Son travail étant destiné avant tout à l'enseignement de l'allemand aux anglophones, Hauch juge bon de réduire la définition précédente pour des raisons pratiques. Il propose alors cette définition:

> *An idiom is defined as an expression that cannot be brought under any recognized syntatical category and, when literally translated into english, conveys no adequate meaning or one essentially different from the actual one.* (85, p. 1)

Cette deuxième définition conduit l'auteur à exclure toute tournure idiomatique allemande ayant un équivalent presque littéral en anglais. *Im Herzen der Stadt* (= au coeur de la ville) est donc exclu de la liste.

Accessibilité

L'introduction donne des renseignements sur la méthode de travail. Elle donne également la liste de tous les textes dépouillés, leur genre, leur origine et la grandeur de chacun d'eux.

Les résultats se présentent sur deux listes. La première est une liste séquentielle (*sequential list*) de 259 tournures idiomatiques, classées d'après leur mot clef et par ordre de fréquence décroissante. Chaque mot clef est précédé d'un numéro de série; il est suivi de l'indice de fréquence et de l'indice de répartition. Ainsi trouvera-t-on au numéro de série 108 de la page 19 les indications suivantes:

> *"108 los was ist --? 15,7"*

Ceci signifie que le mot *los* a été trouvé dans la tournure *was ist los?* et que cette tournure est apparue 15 fois en tout dans 7 textes différents.

Dans la seconde liste, les tournures idiomatiques sont rangées par ordre alphabétique de leur mot clef. Chaque mot clef est donc suivi de la tournure ou des tournures qu'il indexe. Ces tournures sont traduites en anglais. Les trois chiffres qui suivent le mot indiquent respectivement la fréquence, la répartition et le numéro de série dans la première liste.

Donc:

"Angst er hat -- davor, vor ihm, he fears it, him 8; 6;155" (85, p. 44)

signifie que la tournure *er hat Angst davor* se retrouve sous le numéro de série 155 de la première liste et elle apparaît en tout 8 fois dans 6 textes différents.

Sous chaque mot clef, les tournures idiomatiques sont présentées par ordre alphabétique.

2.3.3 Simmons, L.V.T. *A Vocabulary Count Based on Three German Dramas* 1929

But

L'auteur cherche à déterminer l'ordre dans lequel on devrait introduire trois pièces allemandes en classe. Il fait alors l'étude du vocabulaire contenu dans chacune de ces trois pièces.

Méthode

Simmons fait le décompte de tous les mots de chacune des trois pièces: *Minna von Barnhelm* de Lessing, *Wilhelm Tell* et *Jungfrau von Orleans* de Schiller. Il trouve que le nombre de mots différents dans chacune est respectivement 2 820, 3 350 et 3 620. 920 mots sont communs aux trois pièces; 1 120 le sont à *Minna von Barnhelm* et à *Wilhelm Tell*, 1 200 à *Minna von Barnhelm* et *Jungfrau von Orleans* puis 1 520 à *Wilhelm Tell* et *Jungfrau van Orleans*. La corrélation élevée entre les pièces du même auteur est évidente.

2.3.4 *New York State Syllabus* 1930

Un comité fait la révision d'un syllabus préparé en 1910 et contenant entre autres une liste normalisée des mots et tournures idiomatiques de chacune des quatre langues étrangères - français, allemand, italien et espagnol - enseignées dans l'État de New York.

Les mots dans la liste allemande sont groupés par centre d'intérêt et sont au nombre de 1 050. La liste contient 300 adjectifs, 200 verbes irréguliers, 450 verbes réguliers et 100 adverbes. Les conjonctions et les prépositions n'y sont pas.

Le comité prépare deux autres listes, car il prévoyait un vocabulaire pour les trois premières années d'apprentissage. Le vocabulaire total ainsi sélectionné comprend environ 2 200 mots auxquels s'ajoutent quelques 75 tournures idiomatiques.

2.3.5 Griesbach, H. et Schulz, D. *Glossar zu Deutsch Sprachlehre für Ausländer* 1958

But

Le livre se destine à l'enseignement et à l'apprentissage de l'allemand comme langue étrangère.

Méthode de sélection

Les auteurs ne donnent aucune indication sur leur critère ni sur leur méthode de sélection. On peut seulement supposer qu'ils ont choisi les mots nouveaux qui apparaissent dans les différents paragraphes de leur grammaire allemande.

Accessibilité

Il s'agit d'une simple liste de mots qui doit aider l'étudiant étranger à vérifier s'il a bien compris le sens des mots qui lui auront été expliqués en classe. La liste est censée accompagner l'utilisation du livre *Deutsche Sprachlehre für Ausländer, Grundstufe*, et rassurer le débutant dans la révision de ses leçons. Le livre fait partie d'une série de publications bilingue, trilingue, voir même quatrilingue.

L'utilisation de la liste est relativement simple puisque le vocabulaire est groupé d'après l'ordre des paragraphes du livre utilisé en classe.

2.3.6 Michéa, R. *Vocabulaire progressif allemand* 1959

But

L'auteur vise l'enseignement et l'apprentissage de la langue allemande. Il dit dans sa préface: "cet ouvrage... est destiné essentiellement à affermir et développer la connaissance active de l'allemand".

Critères de sélection

Michéa choisit le critère de la fréquence. En effet, ses 1 600 mots fondamentaux suivent un classement alphabétique et statistique (p. 10-112). Mais il tient également compte de la disponibilité; son vocabulaire complémentaire est classé par centres d'intérêt (p. 114-258).

Corpus

Le vocabulaire sélectionné se limite à environ 5 000 unités dont 1 600 mots fondamentaux, 2 800 mots complémentaires; le reste constituant le vocabulaire scientifique.

Méthode

Michéa n'a pas décrit le travail de sélection dans la préface de son livre. Il nous a livré seulement les résultats. Mais, il nous laisse assez bien entrevoir sa méthode de travail. Il s'agissait, pour lui, de déterminer un vocabulaire limité, "formé principalement de mots de sens très large, remployables dans un grand nombre de sujets et de situations et dont la fréquence est effectivement attestée par la statistique..." (p. 3). Sans que l'auteur le dise explicitement, cette conception du mot de sens très large semble se rapprocher de la notion de valence (voir infra 1.4.2.5).

Comme on peut bien s'en douter, l'initiateur du critère de disponibilité ne pouvait pas se satisfaire de la seule étude de fréquence. Le vocabulaire fréquent est suffisant en maintes circonstances mais il reste impropre à rendre les détails et les nuances. Michéa relie donc à ses mots fréquents un vocabulaire complémentaire qui permet un plus haut degré de précision dans la désignation des choses et des êtres, ainsi que dans l'expression de l'action, des qualités et des rapports (voir exemples ci-après). C'est ce vocabulaire qui correspond au vocabulaire disponible.

Michéa distingue un troisième niveau de mots qu'il classe sous la rubrique 'vocabulaire scientifique'. Il s'agit là, certes, aussi de mots disponibles, mais très spécialisés.

Les mots appartenant à ces trois différentes catégories - vocabulaire fondamental, vocabulaire complémentaire et vocabulaire scientifique - ne peuvent pas avoir un même indice de fréquence, car à toute progression dans le sens de la spécification, de la caractérisation, de la finesse correspond une régression dans l'emploi des mots fréquents. Michéa illustre cet aspect par les exemples suivants:

a) *Ein Beamter gab ein Zeichen, und der Zug setzte sich in Bewegung.* (p. 9)

b) *Der Bahnhofsvorsteher gab das Zeichen zur Abfahrt, und der Schnellzug setzte sich in Bewegung.* (p. 113)

c) *Die Rakete wird mit einem Gemisch von Alkohol und fluessigem Sauerstoff betrieben.*

Les phrases a) et b) décrivent la même réalité et ont le même sens. Mais la seconde qui utilise le vocabulaire complémentaire est plus précise que la première qui se limite au vocabulaire fondamental. En effet *der Bahnhofsvorsteher, das Zeichen, zur Abfahrt, der Schnellzug* sont respectivement plus précis que *ein Beamter, ein Zeichen, der Zug.* Mais la proportion des mots fréquents dans la phrase b) baisse. Quant à la phrase c) on ne l'entendra souvent que d'un spécialiste ou d'un ingénieur des fusées.

Cette distinction de trois niveaux de vocabulaire a guidé la présentation des résultats.

Présentation

Le vocabulaire fondamental contient 1 600 mots rangés par ordre alphabétique. L'auteur distingue entre les mots très fréquents, soit 1 000 environ. On se demande ici ce que sont les autres 600 mots, sont-ce aussi des mots fréquents? Sinon, comment l'auteur les a-t-il mis dans ce groupe? Ces mille mots doivent représenter environ 70% des mots nécessaires pour une rédaction en langue allemande.

Le vocabulaire complémentaire est réparti en soixante-cinq chapitres. Un rappel des mots fondamentaux termine chaque chapitre. Ceci est un principe important en didactique des langues. Il s'agit notamment ici de la répétition qui sert à fixer les éléments appris. L'auteur applique une de ses idées qu'il nous présente en 1950 dans son article *La culture par la langue.* (144, p. 130)

> "Une pédagogie bien comprise doit corriger les inévitables imperfections des listes de fréquence, et le grand principe reste: répéter, répéter, répéter. Tout l'enseignement de début consiste à organiser méthodiquement la répétition, active et passive."

Michéa, qui est un professeur expérimenté ne croit certainement pas au simple apprentissage des listes de mots, aussi prévoit-il des locutions, des phrases-types mais parfois aussi de séries d'actions dans lesquelles les mots nouveaux se retrouvent. Le vocabulaire complémentaire se subdivise en trois parties dont la première inclut un vocabu-

laire usuel concret et les deux autres un vocabulaire servant à traduire des manières personnelles de juger, de sentir et de raisonner.

Le vocabulaire scientifique comprend d'abord un vocabulaire général rangé par ordre alphabétique, puis des chapitres élémentaires correspondant aux grandes disciplines scientifiques et à quelques applications courantes.

Quelques pages finales sous la rubrique 'Exploration et mobilisation du vocabulaire. La rédaction' sont destinées à faciliter la recherche du vocabulaire utilisable dans le cadre d'un sujet ou d'une situation donnés.

L'auteur fournit la traduction française des mots, des locutions et des phrases types.

Le génitif et le pluriel des noms sont indiqués pour les substantifs assez difficiles. Le génitif singulier n'est indiqué ni pour les féminins (invariables au singulier) ni pour les masculins et les neutres de la déclinaison forte (leur terminaison étant (- e s). On verra donc à la page 71 par exemple:

a) *das Paar (-e)* - seule la terminaison du pluriel est indiquée;

b) *der Pfarrer (-)* - le trait indique que le mot reste invariable au pluriel;

c) *das Ohr (-s, -en)* - il s'agit ici d'un substantif qui suit la déclinaison mixte.

Un trait d'union entre une particule et un verbe indique que la particule est séparable. C'est le cas dans *heim-kehren, heim-kommen* (p. 49) alors que la particule *ver* dans *versprechen* est inséparable (p. 100).

Utilisation

L'ouvrage de Michéa est assez répandu dans les écoles et universités françaises. C'est un instrument de travail très important aussi bien à l'étudiant qu'au maître. Il servira aussi bien à l'étude générale que spécialisée de la langue allemande. Les écoles de traducteurs et d'interprètes, les ingénieurs et techniciens en formation trouveront en ce livre un document fort utile.

2.3.7 Uhlig, W., Chatelanat, C. et Lang, J.B. *Vocabulaire de base allemand-français* 1963

But

Leur vocabulaire est destiné à l'apprentissage de l'allemand. Les auteurs n'entendent cependant pas encourager l'apprentissage des listes de mots. Mais ils estiment que c'est à l'occasion de telle lecture, de tel exercice de grammaire que le professeur pourra intégrer les mots nouveaux à ceux qui sont déjà connus.

Critère de sélection

Les auteurs ne sont pas explicites à ce sujet, mais ils laissent entendre qu'ils ont tenu compte de la fréquence des mots et notamment des premiers 1 200 mots:

> "Le choix des mots, on s'en doute, n'a pas toujours été aisé. Si l'on dépasse notablement le premier fonds de 1 000 à 1 200 mots les plus fréquents, on atteint une zone où la fréquence d'emploi joue un rôle de moins en moins important, et où il faut recourir à d'autres critères." (Introduction)

Les auteurs ne nous indiquent pas ces autres critères.

Corpus

La liste contient environ 2 500 mots dont 1 500 sont étudiés dans le cours élémentaire d'allemand *Wir sprechen Deutsch* (tome I à III).

Accessibilité

Le vocabulaire est groupé en 18 centres d'intérêt. Les mots présentant une particularité d'emploi ou de sens sont accompagnés d'un exemple. Deux index - dont un en allemand et l'autre en français - terminent le livre.

2.3.8 Wängler, H.H. *Rangwörterbuch* 1963

But

Il s'agit ici d'un travail phonétique dont le but est de donner une vue sur les formes fréquentes qui apparaissent dans le haut allemand de nos jours. La recherche veut donc se consacrer avant tout à la langue parlée. Mais l'auteur entrevoit aussi l'utilisation des résultats dans

l'enseignement de l'allemand comme langue seconde et il espère inciter à la réalisation d'un inventaire plus complet de la *Sprechwirklichkeit* (la réalité du parler...) qu'il considère comme fondement de la langue au sens le plus large.

Critère

Le critère qu'il a choisi est la fréquence. Pour les mots de ses deux sous-ensembles, il ne retient que ceux qui ont une fréquence égale ou supérieure à 10. Pour toute la population lexicale il ne considère que ceux ayant une fréquence supérieure ou égale à 15.

Corpus

Des 160 532 mots courants, Wängler retient 1 022 mots différents dont l'indice de fréquence varie de 4 590 (l'article *die*) à 15 (le substantif *Zukunft*).

Des 80 266 mots courants de la langue parlée l'auteur trouve 735 mots différents dont l'indice de fréquence va de 2 890 (*ich*) à 10 (*zwölf*).

Des 80 266 mots courants écrits il relève 742 mots différents dont la fréquence varie de 2 999 (*der*) à 10 (*übrigens*).

Méthode

Wängler travaille sur des enregistrements et sur des textes écrits. Il enregistre plusieurs personnes de différents âges, surtout celles du nord de l'Allemange à leur insu. Les allemands du nord (région de Hannovre) sont censés parler le plus bel allemand, ou plutôt un allemand qui s'approche le plus du haut allemand. Les enregistrements sont transcrits et dépouillés. Mais l'auteur prend aussi des mots de différentes revues, de journaux quotidiens et des périodiques. Ces mots sont aussi mis sur fiches, ordonnés et transcrits sous la forme dans laquelle l'auteur les entendait le plus fréquemment.

Problèmes

Wängler ne nous donne pas sa définition de l'unité lexicale de décompte. Mais ses résultats nous laissent supposer qu'il entend par mot, une entité de discours avec une certaine forme sonore; aussi les formes *Tages*, *Tagen*, *Tag*, *Tage* sont-elles des mots différents dans sa liste. Sa conception du mot se rapproche donc de celle de Kaeding (p. 24 infra) pour qui les variantes morphologiques sont aussi des mots différents.

La liste contient les noms des villes et ceux des personnes. Pour un étranger qui étudie l'allemand ce ne sont peut-être pas les choses les plus utiles à étudier. Les noms des jours de la semaine sont inclus dans la liste sauf le jeudi. On se demande si dans pareil cas, l'auteur ne devrait pas l'ajouter.

Accessibilité

L'auteur donne d'abord un aperçu des phonèmes allemands en indiquant leur nature et en les accompagnant d'exemples.

Le premier tableau présente les mots par ordre alphabétique. Chacun est suivi de sa prononciation en alphabet phonétique international, puis de trois indices de fréquence dont le premier (U) est la fréquence du mot en langue parlée (enquête sur les enregistrements), le second (Z) est la fréquence du mot écrit et le troisième (G), l'indice de fréquence totale, c'est-à-dire la somme U + Z.

Le deuxième tableau donne la liste de tous les mots par ordre de fréquence décroissante (G). Chaque mot différent a un numéro de série; ce qui fait que les mots de même fréquence ont des numéros de série différents. Ces derniers sont alors rangés par ordre alphabétique. C'est le cas pour les mots des pages 47 - 48 (253) de fréquence 18 et dont les numéros de série vont de 864 (*Arbeiter*) à 910 (*zuviel*).

Le troisième et le quatrième tableaux donnent respectivement les mots de la langue parlée et de la langue écrite, rangés par ordre de fréquence décroissante et précédés d'un numéro de série.

Le livre est facile à consulter et peut être utile à celui qui veut apprendre en même temps le mot et sa prononciation correcte.

2.3.9 Loriche, R. *Le mot et l'idée* 1965

But

L'auteur poursuit deux buts en présentant son vocabulaire: Il vise d'abord à donner une liste de mots pouvant servir à la révision et au contrôle des connaissances; puis des mots dans un texte où ils gardent leur rôle essentiel de véhicule d'idée.

Critère

Aucune indication n'est donnée quant à la façon dont le choix des mots a été fait.

Accessibilité

Les mots sont répartis par centre d'intérêt. L'auteur en trouve 35. Le livre semble plutôt convenir aux étudiants d'une école de traducteurs et d'interprètes ou pour des étudiants avancés qui veulent se spécialiser

2.3.10 Melzer, Friso *Das Wort in den Wörtern* 1965

Il s'agit d'un dictionnaire en 372 pages qui contient environ 140 articles sur le vocabulaire personnel de *Christus-Nachfolge*. Les articles sont classés d'après un point de vue historique. On établit un rapport entre chaque mot et son équivalent en langue anglaise qui est considérée comme langue de l'oecuménisme.

2.3.11 Slaby, H. *Deutscher Wortschatz in Sachgebieten* 1965

But

Le travail de Slaby vise l'enseignement du vocabulaire aux étrangers, mais aussi aux autochtones des cours élémentaires et moyens des lycées (*höhere Schule*).

Méthode

Les mots sont groupés par centre d'intérêt (*Sachgebiet*). L'auteur en distingue dix dont chacun a ses subdivisions. Le centre d'intérêt sur l'homme se subdivise ainsi en dix sections: le corps - le développement - les soins - la maladie - les repas - la nourriture, les boissons - les habits - les parures - la famille: a) membres de la famille, b) événement dans la famille -.

Pour chaque centre d'intérêt l'auteur relève les mots clefs et donne une liste d'autres mots se rapportant à ce mot clef. On revoit ici la notion de champ sémantique comme chez Hauch.

2.3.12 Wolf, F. *10 000 Wörter - A German Vocabulary for Students* 1969

But

L'auteur vise l'enseignement et/ou l'apprentissage de l'allemand et surtout du vocabulaire par centres d'intérêt.

Critère de sélection

La préface signale que les mots ont été choisis selon deux critères: la fréquence et l'applicabilité dans tous les secteurs de la vie quotidienne. L'auteur n'explique pas ce qu'il entend par 'applicabilité'; peut-être voudrait-il exprimer par ce terme, le fait qu'un mot puisse se prêter à l'utilisation et se présenter immédiatement à l'esprit au moment où l'on en a besoin. Dans ce cas applicabilité serait l'équivalent de disponibilité (1.4.2.4).

> *"The vocabulary.....was selected according to practical importance: Frequency and applicability in all sectors of every-day life were decisive."*

Corpus

Le livre présente un vocabulaire d'environ 10 000 mots surtout des substantifs, des verbes et des adjectifs.

Méthode

Aucune indication ne nous permet de découvrir la méthode de travail. L'ouvrage se présente comme une étude de la disponibilité des mots allemands. L'auteur distingue trente centres d'intérêt.

Problèmes

On peut se demander si le choix des mots s'est fait d'après des critères objectifs ou plutôt subjectifs. La fréquence dont parle l'auteur représente-t-elle la fréquence de chacun des mots dans l'ensemble du corpus ou plutôt dans l'ensemble des textes se rapportant à un centre d'intérêt donné.

La liste ne donne aucun mot grammatical comme préposition, pronom, adverbe, mais plutôt des mots thématiques (1.4.2.4): substantifs, verbes et adjectifs.

Accessibilité

Le vocabulaire est classé par centre d'intérêt. Chaque substantif est suivi d'une indication de son genre et de sa forme de pluriel, mais aussi d'une traduction en anglais:

Centre d'intérêt: La ville

Stadt f ("e) *town, city* (270, p. 75)

f = féminin

("e) = désinences du pluriel *die Städte* = les villes.

A la suite d'un mot clef viennent d'autres mots qui recouvrent le même champ sémantique ou qui s'en rapprochent. Pour le mot clef *Stadt*, l'auteur donne 10 autres substantifs:

Grossstadt	*Vorort*
Kleinstadt	*Umgebung*
Weltstadt	*Einwohner*
Stadtviertel	*Adressbuch*
Vorstadt	*Städter*
1 adjectif:	*benachbart*
1 verbe:	*umgeben*

1 tournure idiomatique: *am Rande der Stadt.*

Cette façon de présenter ces mots nous permet de croire que Wolf a fait une étude de la disponibilité. Car tous les mots qu'il donne sont susceptibles de venir à l'esprit quand on se met à penser à la ville.

En appendice l'auteur donne d'une part la transcription phonétique des mots difficiles et d'autre part les infinitifs et les temps primitifs (présent, imparfait, participe passé) des verbes irréguliers apparaissant dans sa liste.

Les étudiants des écoles de traduction ou d'interprètes (anglais-allemand) trouveraient en ce livre un bon manuel pour l'apprentissage systématique du vocabulaire. Les professeurs d'allemand spécialisé, allemand scientifique par exemple, peuvent aussi en tirer profit.

2.3.13 Schmidt, P. *Der Wortschatz von Goethes "Iphigenie* 1970

But

L'auteur veut promouvoir l'analyse du poète allemand Geothe. Son travail doit représenter un document pour d'autres recherches sur la pièce "Iphigenie" et être une introduction à la nouvelle méthode de recherche littéraire.

Corpus

Le travail de décompte se fait sur les deux versions de la pièce. Il s'agit d'un texte en prose de 17 274 mots et du même texte en vers de 14 568 mots.

Méthode et problèmes

Le décompte a été fait par des machines. L'auteur définit le mot comme une suite d'éléments compris entre deux vides. *die Elementenfolge zwischen zwei Leerstellen.* Cette définition est pratique et compréhensible pour une machine, mais elle entraîne un certain nombre de problèmes. En effet la machine, qui est très docile, fonctionnait en tenant compte des apostrophes et des traits d'union. Les mots qui en avaient étaient alors divisés en mots différents. Il apparaît alors des mots allemands n'ayant pas de sens tel que le pronom *es* dans sa forme *'s*; *heilgem* devient deux mots *heil* et *gem* (217, p. 35); *gem* n'a pas de sens.

Accessibilité

Peter Schmidt donne un tableau de tous les mots qui reviennent dix fois ou plus dans la pièce.

Cette étude fait partie d'une série *Indices zur deutschen Literatur* qui est une application de la lexicométrie aux études littéraires notamment à l'analyse des poètes et écrivains.

2.3.14 Kopps, H. et Münster, T. *Gesamtverzeichnis der Verben aus dem Corpus zum "Wissenschaftsdeutsch"* 1971 (manuscript non publié)

Le travail donne une liste de tous les verbes apparus dans un corpus de textes scientifiques lors d'une étude de la fréquence des mots. Les auteurs ont préparé cette section comme fondement à une étude de la formation des mots allemands (*Wortbildung*). Pour l'instant la liste est un document interne à l'usage de *Arbeitsstelle für wissenschaftliche Didaktik* à Munich.

Corpus

L'analyse a porté sur trois genres de textes scientifiques: *Naturwissenschaft NW* (sciences naturelles), *Sozialwissenschaft SW* (sociologie), et *Geisteswissenschaft GW* (les lettres) qui avaient en tout 3 315 verbes différents revenant 43 000 fois.

Problèmes

Les auteurs tiennent uniquement compte de la fréquence. On peut se demander si le critère répartition n'aurait pas ajouté quelques précisions surtout si on envisageait d'utiliser cette liste pour le choix des verbes les plus utiles. En se décidant de sélectionner tous les verbes d'indice de fréquence 10 ou plus par exemple, les verbes *dezentralisieren, konvergieren, verbrennen* et *objektivieren* seront pris mais pas *dominieren* ni *überordnen*. Or, l'examen de la liste nous montre que ces deux derniers verbes sont beaucoup mieux répartis tel que l'indique le tableau suivant:

Verbe	fréquence	fréquence dans les catégories de textes		
		NW	*SW*	*GW*
dezentralisieren	13	0	13	0
konvergieren	20	20	0	0
verbrennen	16	16	0	0
objektivieren	17	0	0	17
dominieren	9	3	4	2
überordnen	9	5	2	2

Les quatre premiers infinitifs apparaissent plutôt comme des verbes uniquement disponibles. Pour qui devrait lire les textes de science naturelle le verbe *konvergieren* serait très utile mais en est-il de même pour celui qui aurait à lire des textes se rapportant aux deux autres catégories? Par contre les verbes *dominieren* et *überordnen* sont utiles pour un lecteur qui travaille sur les trois genres de textes.

Accessibilité

Le premier tableau donne la liste de tous les verbes en indiquant la fréquence totale puis la fréquence dans chacun des trois genres de textes d'après l'exemple suivant:

	fréquence totale	1er genre	2ème genre	3ème genre
sein	5 697	1 975	1 764	1 958

Le deuxième tableau donne la liste des verbes par ordre de fréquence décroissante.

Le travail de ce genre est sans aucun doute utile pour l'enseignement et l'apprentissage de l'allemand scientifique.

2.3.15 Steger, H. et Keil, M. *Wortliste für VHS-Zertifikat* 1971

But

Il s'agit d'une liste de mots qui devront être connus par un élève qui postule le grade de Certificat d'allemand comme langue étrangère. Les auteurs précisent que ce vocabulaire représente un minimum que les auteurs de manuels auront à compléter. La liste se veut être utile aussi bien à l'étranger dans une zone germanique qu'à l'autochtone.

Corpus

La liste finale apporte environ 2 000 mots.

Critère de sélection et méthode

Les auteurs font remarquer que le critère de disponibilité n'a pas éliminé les faiblesses des études de fréquence. Notre analyse (1.4.2.1) nous montre qu'ils ont raison jusqu'à un certain point, mais ils passent sous silence les autres critères objectifs qui peuvent raffiner la méthode de sélection du vocabulaire de base.

Pour choisir leurs mots un groupe de sept théoriciens et praticiens s'est donné comme tâche de sélectionner individuellement un nombre de mots ne dépassant pas 2 000. En réunissant leurs 7 listes individuelles ils obtiennent une liste commune d'environ 4 000 mots. De ce total ils retiennent seulement ceux qui avaient été choisis par plus de la moitié du groupe. Ceci leur donne une liste fondamentale de 1 200 mots. Il s'agit donc ici d'un choix de mots empirique ou peut-être 'au hasard'. Comme les auteurs voulaient atteindre un maximum de 2 000 mots, ils décident de compléter leur liste fondamentale avec des mots pouvant introduire une conversation:

> *"Es schien uns wichtiger, Hilfsmittel für die Anknüpfung von Gesprächen zu geben, statt lediglich Sachbenennungen vorzustellen."* (233, p. IV)
>
> (Il nous a paru plus important de donner des instruments pour déclencher des conversations au lieu de présenter de simples termes d'identification.)

Dans ce but ils choisissent alors des exemples de phrases et de syntagmes. Afin d'ordonner et de classer le matériel lexical et syntaxique, les auteurs devaient faire face à certains problèmes: l'un des plus importants était de trouver quelles sont les habitudes et les connaissances linguistiques d'un citadin moyen? Pour y répondre une commission fait une enquête originale quelque peu semblable à une étude de la disponibilité dont les centres d'intérêts s'appellent *Redekonstellation*. Au lieu de dresser une liste de mots qui viennent à l'esprit lorsqu'on pense à un centre d'intérêt quelconque, on demande à des personnes d'improviser une scène dans le cadre d'un thème qui leur est donné. Ces 'acteurs' peuvent improviser la conversation d'un hôte et d'un garçon de restaurant, la conversation d'un médecin et de son patient. D'une telle conversation on tire un vocabulaire qui peut plus vite conduire à la conversation que les mots disponibles choisis d'après la méthode du *français fondamental*. En effet ces derniers peuvent plutôt aider à faire une description ou narration qu'une conversation.

Steger et Keil semblent alors avoir introduit un nouvel aspect à l'étude de la disponibilité avec leur méthode de la *Redekonstellation*.

Accessibilité

Toutes les unités lexicales sont classées par ordre alphabétique. Les verbes sont suivis de leurs temps primitifs.

Dans de courtes phrases on donne aux débutants les sens les plus utiles d'un mot.

Un premier appendice présente les termes mathématiques de tout le vocabulaire.

Le second appendice relève les préfixes ainsi que les verbes avec lesquels ils peuvent se construire.

2.4 Pfeffer, J.A.

La section précédente nous a présenté des auteurs qui ont travaillé assez indépendamment du *Dictionnaire de fréquence*. Les résultats de la recherche de Kaeding n'ont pas été utilisés pour dresser leurs listes. Dans les lignes qui suivent nous verrons un auteur original contemporain. Ses publications sur la lexicométrie allemande commencent en 1962 et ne cessent d'augmenter. Si on les classait toutes d'après notre système on verrait que l'une d'elles *Basic (spoken) German word List, Mittelstufe* (191) serait au nombre des travaux basés sur Kaeding. En effet, Pfeffer choisit des termes lexicaux de plus haute fréquence et répartition dans sept travaux ultérieurs basés sur les résultats de Kaeding. Mais toutes les autres publications appartiennent au groupe 'autres travaux'.

En parlant de Pfeffer nous mettrons l'accent surtout sur son travail original sur le *Grunddeutsch*.

C'est en 1960 que commence le travail sur l'allemand fondamental à Munich. Il est alors subventionné et encouragé par le *Language Development Section of the U.S. Office of Education*. Pfeffer s'inspire des travaux français et espagnols pour bâtir son *Grunddeutsch*. Parlant du *Vocabulario español* et du français fondamental il dit notamment:

> *"... indeed, they have pointed the way to additional refinements in compiling word lists and have served as guideposts in the evolution of Grunddeutsch, of which the Grundstufe (level one) of the Basic (spoken) German Word List is the first part."* (187, p. 7)
>
> *Basic (spoken) German Word List, Grundstufe*, 1964, (187)

But

Dans ce premier volume du *Grunddeutsch*, l'auteur vise l'enseignement et l'apprentissage de l'allemand parlé contemporain.

Critère

Quatre critères sont à la base de la sélection des mots chez Pfeffer: la fréquence, la répartition, la disponibilité et l'empirisme. Ils ont été utilisés dans cet ordre et ont déterminé la méthode de son travail.

Corpus

Afin de compenser la diversité des dialectes allemands, il est suggéré de collectionner un corpus de 600 000 occurrences de sources orales dans diverses régions d'Allemagne, d'Autriche et de Suisse; puis 833 000 mots disponibles sont dressés spontanément par des élèves. De ces deux grands totaux, Pfeffer obtient une liste fondamentale de 1 084 mots auxquels il ajoute 185 autres pour compléter son *Grunddeutsch, Grundstufe* de 1 269 mots.

Méthode

Un des soucis de Pfeffer était que le gros travail matériel soit la base de nombreuses études complémentaires et ne se limite pas au seul décompte statistique. Aussi a-t-il pris de nombreuses précautions et fait quelques expériences préliminaires. Nous ne relevons ici qu'un certain nombre de ces précautions qui peuvent être utiles pour un initié:

Le nombre d'enregistrements et la durée de chacun: Ayant prévu un total d'éléments à collectionner on peut se faire une idée de la quantité de bande à utiliser. Pfeffer en utilise plus de 400 et chacun de ses enregistrements dure environ 12 minutes.

Zones linguistiques: Un bon choix est nécessaire pour avoir un échantillon représentatif. Pfeffer choisit 70 villes et villages dans la zone germanique. Il détermine la population de chaque lieu, la vocation, l'éducation, l'âge et le sexe de chaque sujet interviewé, la grandeur et le lieu de résidence de ce dernier pendant ses années de formation.

Travail pratique

Pour les enregistrements il se fait assister par des spécialistes du son et par des linguistes.

Les bandes originales sont conservées aux archives et le dépouillement ne se fait qu'avec des copies. Lors des codages il inclut les numéros de l'enquêteur et du sujet.

L'équipe d'enregistrement compte 57 enquêteurs dont 8 ont été interviewés. 25 thèmes couvrant la répartition des intérêts quotidiens de l'homme avaient été pré-sélectionnés. La conversation sur ces thèmes constituait une bande enregistrée. Les dialogues ayant plus de 10% de tournures dialectales étaient rejetés. Après cette élimination il fut retenu 401 enregistrements sur 409. Les enregistrements furent transcrits. Une tabulation des 595 000 cartes donna un total d'environ 25 000 unités lexicales dont chacune représente la somme de ses propres formes fléchies d'après l'exemple suivant:

887 *Wissen*

208 *wissen*

539 *weiss*

23 *weisst*

1 *wisst*

3 *wusst'*

55 *wusste*

25 *wussten*

8 *wüsste*

1 *wüssten*

1 *wissenden*

23 *gewusst*

De ces 25 000 unités lexicales on exclut les noms propres et de lieux, ainsi que les adjectifs qui en dérivent (sauf *deutsch* = allemand).

Les mots les plus communs sont déterminés en tenant compte de la fréquence et de la répartition dont les indices minima sont respectivement 40 et 25. La répartition valait la peine d'être mesurée ici étant donné que le nombre d'enregistrements total s'élevait à 450 - soit 401 interviews + la voix des 57 enquêteurs moins les 8 enquêteurs interviewés. Combinant les deux critères à une analyse empirique des mots quant à leur applicabilité, universalité et indispensabilité, l'auteur réduit son total à 737 mots de la langue parlée et voilà une première phase de son travail.

Pour la deuxième phase, l'étude de la disponibilité, le professeur Pfeffer prévoit 21 centres d'intérêt. Des élèves de 15 - 16 ans trouvent spontanément des substantifs, des verbes et des adjectifs se rapportant à ces centres. Le codage est fait pour un pays, l'état et la ville. On donne aussi les indications suivantes: l'école, la classe, le sexe, vocation du père, origine des parents, lieu et grandeur de la localité de résidence pendant les années de formation. Tous ces détails que le professeur Pfeffer s'est arrangé à prendre peuvent être fort utiles à d'autres recherches psychologiques, psycho-pédagogiques et psycholinguistiques et même socio-linguistiques. Les tests de disponibilité ont été donnés dans 82 écoles de 48 villes différentes. L'âge de 15 - 16 ans a été choisit (et non 9 - 12 ans) après les expériences préliminaires qui ont fait constater que le vocabulaire spontané n'était plus enfantin et pas encore sophistiqué. L'enquête fournit 833 000 mots au total représentant 34 000 mots différents. Leurs indices de fréquence varient entre 960 et 1. Des 34 000 mots seulement 347 sont retenus pour la liste finale. On tient compte de la grande répartition de l'unité lexicale et de son indice de fréquence qui devait être supérieur à 100 pour la retenir. Après cette deuxième phase la liste a désormais 1 084 mots.

La dernière phase du travail est celle de la collection empirique de 185 mots. Ces derniers sont soigneusement sélectionnés et représentent:

a) des parties d'une séquence directe, comme *sechzehn* (seize) et *sechzig* (soixante) qui se rapportent à *sechs* (six) dont la fréquence est de 190 et l'indice de répartition 25. De même *gestern* (hier) et *morgen* (demain) par rapport à *heute* (aujourd'hui) (813/255);

b) des parties d'une séquence associative comme *Mond* (lune), *Stern* (étoile) par rapport à *Sonne* (soleil) (50/36);

c) des termes liant une chose spécifique à un tout ou un tout à une chose spécifique: *Hund* (chien) 731 et *Katze* (chat) 708 par rapport à *Tier* (animal); ou *Mund* (bouche) 553 et *Nase* (nez) à *Gesicht* (visage);

d) des antonymes manquant: *drinnen* (à l'intérieur, dedans) complétant *draussen* (dehors) 135/81, ou *Antwort* (réponse) complément de *Frage* (question) 257/215;

e) des parties d'un composé commun comme *Rad* (roue) dérivé de *Fahrrad* (bicyclette) 375, ou *Motor* dérivé de *Motorrad* (motocyclette) 292;

f) des groupes topiques comme les noms de mois, des poids et mesure;

g) des notions comme *taub* (sourd) et *stumm* (muet) etc.

Cette partie empirique consiste, en gros, à analyser les résultats des enquêtes statistiques et d'y ajouter ce que les statisques n'arrivent pas à trouver. Des 185 mots additionnels les trois quarts étaient déjà inclus directement ou indirectement soit dans la liste des mots disponibles, soit dans le corpus de la langue parlée. C'est le dernier quart qui a effectivement été glané empiriquement.

Problèmes

Dans son volume I, *Grundstufe*, Pfeffer considère comme unité de décompte la forme du mot; mais il a aussi eu à résoudre le problème de la définition du mot dans la suite. Sa conception se rapproche de celles de Bloomfield (*Language* 1933): *a word is a minimum free form* et de Ullmann: *ein Wort ist morphologisch ein ein- oder mehrteilig freies Mindestgebilde.*

En effet la définition du mot chez Pfeffer se retrouve dans celle qu'il donne de tournure idiomatique. Il considère que les champs sémantiques d'un mot se divisent en quatre groupes dont un est libre, un autre partiellement libre - un troisième partiellement limité *teilweise begrenzt - partially restricted* et le dernier est limité *begrenzt - restricted.* Le mot est libre dans sa signification fondamentale comme par exemple "tomber, dans (la pomme tombe) *der Apfel fällt.* Le mot est partiellement libre au sens figuré (*übertragene Bedeutung*) comme (tomber) *fallen* dans *die Preise fallen* (les prix baissent), *das Fieber fallt* (la fièvre tombe). Le mot est partiellement limité si son sens est conditionné par une liaison avec un autre mot interchangeable: c'est comme

le mot *halten* dans *eine Rede halten* est conditionné par *Rede* (discours) que l'on peut remplacer par *Predigt*, *Vortrag* etc. Un mot est limité si son sens primordial se transforme lorsqu'il est en liaison avec d'autres mots. C'est le cas de *fallen* avec *flach* dans *flach fallen*.

Ces définitions jouent un grand rôle dans les recherches ultérieures de Pfeffer, notamment dans la classification sémantique des mots. Dans le volume II de la série *Grunddeutsch*, on a relevé les 75% des mots libres et partiellement libres les plus utiles. Les autres mots limités ou partiellement limités, c'est à dire au fond les mots idiomatiques, ont été réservés pour le volume III de la série.

Dans l'étude de la disponibilité en allemand, un problème se pose généralement. C'est celui des noms substantivés. Tout infinitif allemand peut être substantivé. Un élève peut donc, en remplissant sa feuille, inscrire un verbe comme infinitif mais aussi comme substantif et les compter comme deux mots différents. Pour éviter ce problème Pfeffer décide que certaines écoles collectionnent uniquement des substantifs alors que d'autres relèvent les adjectifs et les verbes.

Accessibilité

Le *Basic (spoken) German, Grundstufe* présente trois tableaux. Au premier tableau les mots sont rangés par ordre alphabétique. Le second donne un groupement en partie du discours (les substantifs, les verbes, les adjectifs, les adverbes, les prépositions, les conjonctions et interjections; les articles, les pronoms et les nombres).

Le troisième tableau donne les mots par ordre de fréquence décroissante et par groupe d'origine (orale, topique ou empirique); pour cette dernière origine les mots sont rangés par ordre alphabétique.

Le livre se termine par deux appendices dont le premier donne les mots thématiques et le second les mots régionaux. En dehors de *brausen*, *Fleischer* et *Kusine* (appendice B) les autres mots ont un indice minimum de fréquence de 175.

Innovations

Sur le plan scientifique le vocabulaire de Pfeffer représente pour l'instant une des listes les plus recommandables pour la langue allemande. L'objectivité de la sélection l'emporte sur la subjectivité. La méthode empirique a été réduite au minimum pour céder place à la fréquence, à la répartition et à la disponibilité.

Il pense à la gradation de ses éléments dans l'enseignement et l'apprentissage de l'allemand. Aussi prévoit-il, dans sa recherche, trois niveaux de vocabulaire. Il en était en 1970 au deuxième, *Mittelstufe*.

Par rapport au français fondamental Pfeffer fait d'énormes progrès: Il augmente le nombre de centres d'intérêt, celui des élèves pour faire le test; il inclut aussi les verbes et les adjectifs dans l'étude de la disponibilité; les mots de la langue parlée sont aussi augmentés. L'extension de ces deux sources a pour avantage de réduire au minimum la nécessité de faire des additions empiriques.

2.5 TABLEAU CHRONOLOGIQUE DE LA LEXICOMETRIE ALLEMANDE

Un des buts de cette étude étant de faire un inventaire historique de la lexicométrie allemande, nous avons jugé bon de présenter sous forme de tableau chronologique un certain nombre de travaux lexicométriques allemands. Notre liste n'a pas la prétention d'avoir atteint l'exhaustivité mais espère donner les ouvrages les plus remarquables entre Kaeding et nos jours. Pour chacun nous indiquons l'année de publication, le nom de l'auteur, le titre ou l'objet du travail enfin le lieu de publication ou le nom de l'éditeur. Dans certains cas nous indiquons le nom de la revue où il est apparu. Tous ces ouvrages n'ont pas pu être analysés dans notre étude pour la raison principale qu'ils ne nous ont pas été accessibles ou accessibles à temps, comme le travail de Helmut Meier.

Pour avoir de plus amples informations sur un ouvrage donné, il suffira de consulter la bibliographie de cette étude.

TABLEAU CHRONOLOGIQUE DE LA LEXICOMETRIE ALLEMANDE

#	Année	Auteur	Titre (ou objet de la recherche	Lieu publication (ou éditeur)
			AVANT 1898	
-	1823	Gabelsberger, F.	*Silbenlexicon*	
-	1846	Förstermann, E.	*Numerische Lautverhältnisse*	
-	1869	Meyer, L.	Décompte de mots parlés	Berlin
-	1891	Mohr, J.	*Zählungen zur Lauttreue*	
			1898 - 1910	
1)	1898	Kaeding, F.W.	*Häufigkeitswörterbuch d.d. Sprache*	Berlin
2)	1905	Behaghel, O.	*Zählungen zum Gebrauch des Beiwortes*	(HtM)
3)	1909	Frank, L.	*Statistische Untersuchungen.....*	(HtM)

#	Année	Auteur	Titre (ou objet de recherche)	Lieu de publication (ou éditeur)
			1911 - 1920	
4)	1911	Schlachter	*Statistische Untersuchungen.....*	(HtM)
5)	1912	Frank, E.	*Zählungen u.a. des Dativ-E.....*	
6)	-	Prehn, A.H.	*Practical Guide to German Vocabulary*	Oxford University Press
7)	1914	Meras & Miller	*Ein Wortschatz*	
8)	-	Proksch, A.	*Der Wortschatz Theodor Storms*	
9)	1919	Köster, M.	*Die am häufigsten gebrauchten Stämme.....*	
			1921 - 1930	
10)	1921	Schlag, J.	*Häufigkeitsproben aus dem Sprachschatze*	(HtM)
11)	1927	Thurnau, A.C.	*Vocabulary Building German for Beginners*	New York
12)	1928	Bierwirth, H.C.	*Words of Frequent Occurrence in Ordinary German*	New York
13)	1928	Morgan, B.W.	*German Frequency Word Book*	New York

#	Année	Auteur	Titre (ou objet de recherche)	Lieu de publication (ou éditeur)
14)	-	Wadepuhl, W.	*German Idioms*	GQ
15)	1929	Hauch, E.F.	*German Idiom List*	New York
16)	-	Simmons, L.V.T.	*A Vocabulary Count.....*	MLJ
17)	1930		*New York State Syllabus*	New York

1931 - 1940

18)	1931	Engel, E.F.	*The Use of a Standardized Vocabulary...*	MLJ
19)	-	Meier, H.	*The 1 000 Most Frequent German Words*	New York
20)	-	Purin, C.M.	*A Standard Vocabulary of 2 000 Words...*	Chicago
21)	-	Schinnerer & Wendt	*A Suggested List of 1 000 Active German Words*	GQ
22)	1933	Vail, C.I.	*Basic Word and Idiom List*	Albany, N.Y.
23)	1934	Bakonyi, H.	*Die gebräuchlichsten Wörter der deutschen Sprache*	Munich
24)	-	Hohlfeldt, A.R.	*Minimum Standard Vocabulary for German*	GQ
25)	-	Wadephul & Morgan	*Minimum Standard German Vocabulary*	New York

#	Année	Auteur	Titre (ou objet de recherche	Lieu de publication (ou éditeur)
26)	1935	Ortmann, A.A.	*A Study in First-Year German Vocabulary*	GQ
27)	1937	Purin, C.M.	*A Standard German Vocabulary*	Boston
28)	-	Schneider, J.	*Deutscher Wortschatz.....*	Batavia
29)	1938	Twaddell	Décompte des mots parlés	Kopenhagen
30)	1940	Eaton, H.S.	*Semantic Frequency List*	New York
			1941 - 1950	
31)	1946	Muegge, M.A.	*German Word-Frequency Vocabulary*	Pitman
32)	1949	Michéa, R.	*Sprachtatistik, Unterricht*	Paris
			1951 - 1960	
33)	1951	Jewish Vocational Service	*Words for Work*	Boston
34)	1952	Meir, H.	*Die tausend häufigsten Wortformen der deutschen Sprache*	Mspe
35)	1953	Auber, M.	Guide de Fréquence allemand	Gap:ophrys
36)	1956	Michéa, R.	Vocabulaire allemand d'utilité scolaire et générale	RLV
37)	-	Schultze, W.	*Der Wortschatz in der Grundschule*	(in HtM)

#	Année	Auteur	Titre (ou objet de la recherche	Lieu publication (ou éditeur)
38)	1958	Griesbach & Schulz	*Glossar zu 'Deutsche Sprachlehre'*	Munich
39)	1959	Meldau, R.	*Kleines Deutschlandbuch für Ausländer*	Munich
40)	-	Michéa, R.	Vocabulaire allemand progressif	Paris
41)	1960	Taylor & Gottschalk	*A German-English dictionnary of Idioms*	Munich

1961 - 1970

42)	1961	Zellweger, R.	Le vocabulaire du bachelier	Lausanne
43)	1963	Kaufholz, et al.	Vocabulaire de base de l'allemand	Paris
44)	-	Uhlig, et al.	Vocabulaire de base allemand-français	Lausanne
45)	-	Wängler, H.H.	*Rangwörterbuch hochdeutscher Umgangssprache*	Marburg
46)	-	Pfeffer, A.	*Grunddeutsch (Grundstufe)*	New Jersey
47)	1964	Siliakus, H.J.	*The 1 000 Most Useful Words in German*	Babel
48)	1 -	Loriche	Le mot et l'idée	OPHRYS
49)	-	Melzer, F.	*Das Wort in den Wörtern*	Tubingen
50)	1965	Scherer, G.	*Word Frequency in the Modern German Short-story*	Colorado

#	Année	Auteur	Titre (ou objet de la recherche)	Lieu publication (ou éditeur)
51)	-	Slaby, H.	*Deutscher Wortschatz in Sachgebieten*	Frankfurt/M
52)	1966	Oehler, H.	*Grundwortschatz Deutsch*	Stuttgart
53)	1967	Pfeffer, J.A.	*Basic (spoken) German Idiom List, Level I*	New Jersey
54)	-	Swenson, R.	*A Frequency Count of Contemporary German Vocabulary*	Univ. Minnesota
55)	-	Zimmermann, J.	*Word Frequency in the Modern German...*	GQ
56)	1968	Schanze, H.	*Index zu Novalis 'Heinrich von Ofterdingen'*	Frankfurt/M
57)	1969	Griesbach & Schulz	*1 000 idiomatische Redensarten*	Berlin
58)	-	Oehler, et. al.	*Grundwortschatz.....*	Stuttgart
59)	-	Schanze, H.	*Index zu Heinrich von Kleist*	Frankfurt/M
60)	-	Wolf, F.	*10 000 Wörter. A German Vocabulary for Students*	Munich
61)	1970	Pfeffer, J.A.	*Grunddeutsch - Mittelstufe-*	Pittsburgh
62)	-	Pfeffer, J.A.	*Basic (spoken) German Dictionnary*	New Jersey
63)	1970	Schmidt, P.	*Der Wortschatz von Goethe's 'Iphigenie'*	Frankfurt/M
64)		Wohlgemuth-Berglund	*Wort für Wort*	Munich

#	Année	Auteur	Titre (ou objet de la recherche	Lieu publication (ou éditeur)
			1971 -	
65)	1971	Kopps & Münster	*Gesamtverzeichnis der Verben.....*	Munich
66)	-	Steger & Keil	*Wortliste für das VHS-Zertifikat*	Frankfurt/M

CHAPITRE III

CONCLUSION

3.1 Conclusions théoriques

- 3.1.1 Sur la quantité des travaux allemands
- 3.1.2 Sur la méthode
  - a) Critères
  - b) Procédés
- 3.1.3 Sur les problèmes
  - a) Linguistiques
  - b) Statistiques
  - c) Présentation des résultats
- 3.1.4 Sur l'évolution
  - a) De la méthode
  - b) De l'intérêt pour l'enseignement
  - c) De la mécanisation

3.2 Conclusions pratiques
Utilité de la présente recherche

3.3 Déficiences de l'analyse

- 3.3.1 Travaux manquants
- 3.3.2 Grille d'analyse non complétée

3.4 Suggestions pour des analyses ultérieures

- 3.4.1 Procéder par des questionnaires
- 3.4.2 Autre classement des travaux
  - a) Travaux primaires
  - b) Travaux secondaires
  - c) Travaux mixtes
- 3.4.3 La valence lexicale
  La disponibilité

3.1 CONCLUSIONS THEORIQUES

3.1.1 Sur la quantité des travaux allemands

Nous n'avons analysé que vingt-neuf listes de mots; mais le tableau chronologique nous montre qu'il en est resté presqu'autant. Si on s'en tenait alors au seul facteur quantité de travaux sur le vocabulaire allemand, la présente étude valait bien la peine d'être entreprise. Le nombre élevé de listes dressées après Kaeding pour la seule langue allemande nous pousse à croire que l'estimation de Savard et Richards (211, p. 21) du nombre de listes est assez faible. En effet, l'appendice A nous présente en 1971 plus de trente langues du monde pour lesquelles il existe (ou se prépare) une liste; si pour l'allemand il en existe un minimum de soixante, on devrait pouvoir aussi s'attendre à un tel minimum pour le français, l'anglais et peut-être l'espagnol et le russe. Si tel était le cas nous dirions alors que plus de 450 - au lieu de 80 - listes de mots sont apparues depuis la recherche de Kaeding.

3.1.2 Sur la méthode

a) Les critères de sélection

La fréquence et l'empirisme sont les deux critères de sélection qui dominent les recherches sur l'allemand. Mais on ne peut plus dire que la fréquence soit encore considérée comme Le Critère de sélection. Les lexicométristes germanistes s'en rendent compte. Quant à l'empirisme, il ne devrait pas tarder à avoir une importance très réduite chez tous les chercheurs, comme Pfeffer en donne l'exemple dans son travail. Le travail de Steger et Keil, bien que fondé sur l'empirisme, introduit une autre façon de choisir les mots disponibles pouvant facilement mener l'étudiant à la conversation (2.3.15). Cette notion de *Wortkonstellation* mériterait la peine d'être suivie de près et expérimentée pour d'autres langues.

b) Les procédés

La majorité des travaux s'est faite à la main. Ce sont les chercheurs des années soixante et soixante-dix qui utilisent le plus les machines pour leurs dépouillements de texte.

3.1.3 Sur les problèmes

Nous distinguons trois sortes de problèmes: les problèmes linguistiques, statistiques et les problèmes de présentation.

La question de la définition de l'unité de décompte fait partie des problèmes linguistiques. Les sténographes et les phonéti-

ciens comptent plutôt les formes. C'est le cas de Kaeding et Wängler. D'autres chercheurs essaient d'introduire la notion de champ sémantique et regroupent alors les formes sous un mot clef. Hauch est l'un d'eux. Les lexicométristes modernes utilisant des machines sont presque tenus de considérer comme unité de décompte tout ensemble de caractères compris entre deux espaces. Cette dernière définition est très pratique mais n'épargne pas de gros problèmes qui peuvent être différents d'une langue à une autre. Pour l'allemand, il peut apparaître alors des mots qui n'existent pas ou qui n'ont pas de sens (2.3.13).

Parmi les problèmes statistiques, nous mentionnons celui de l'échantillonnage. Kaeding et Pfeffer ont donné quelques exemples d'échantillonnage. Leurs corpus (textes et langue parlée) sont bien représentatifs. A l'échantillonnage se lient les problèmes de la longueur et de l'âge des textes. Kaeding essaie de choisir des textes de même longueur. Pfeffer choisit l'étude de la langue contemporaine. Il en est de même de l'*Institut für Deutsche Sprache* dont le corpus est basé sur des textes de 1945 à nos jours. Cette règle n'a pas toujours été observée chez les différents lexicométristes.

La présentation des résultats varie aussi chez les auteurs; mais il semble que la majorité tend à donner au moins une liste de mots rangés par ordre de fréquence décroissante, si c'est la fréquence qui est le critère primordial de sélection. Plusieurs auteurs n'indiquent pas leur méthode de travail comme Kaeding, Pfeffer, Steger et Keil le font. S'ils l'indiquaient, cela donnerait peut-être plus de confiance à l'utilisateur de leurs ouvrages.

3.1.4 Sur l'évolution

Même sans avoir analysé tous les travaux de lexicométrie allemande, les quelques listes examinées dans cette étude dénotent une certaine évolution à deux points de vue.

D'abord au niveau des critères de sélection. De la seule fréquence chez Kaeding, on arrive chez Pfeffer avec quatre critères: la fréquence, la répartition, la disponibilité et l'empirisme. Hauch représente une étape intermédiaire en ce sens qu'il tient aussi compte de la fréquence et de la répartition (2.3.2). En plus de cela, Bakonyi (2.2.7) entrevoit aussi la notion de disponibilité.

On remarque ensuite un intérêt grandissant pour l'enseignement des langues. C'est Morgan qui ouvre la porte en transformant les travaux pour sténographie en une liste de mots utilisables dans l'enseignement. L'aspect didactique se fait encore mieux sentir lorsqu'on commence à concevoir des vocabulaires adaptés à des niveaux d'étude; c'est-à-dire la notion de gradation chez Vail 1933, Wadepuhl 1934 puis Pfeffer dans son *Grunddeutsch* qui est conçu à

trois niveaux (*Grundstufe, Mittelstufe et Oberstufe*).

La lexicométrie tend à se mécaniser. La plupart des centres de recherche (Appendice B) disposent des machines ou travaillent en collaboration avec des écoles techniques.

L'examen du tableau chronologique (page 64) montre que les années soixante ont été des plus fructueuses dans les travaux de lexicométrie allemande. Il nous reste à espérer que les années soixante-dix le seront plus, ou du moins autant.

3.2 CONCLUSIONS PRATIQUES

La présente étude servira de document de référence au didacticien, au professeur d'allemand, à ceux qui ont à élaborer ou à choisir des méthodes d'enseignement, au chercheur et à l'étudiant. L'analyse des travaux donnera un premier contact avec un ouvrage qu'on veut utiliser. L'auteur des méthodes se choisira un vocabulaire de base sur lequel fonder son enseignement. L'inspecteur d'enseignement fera également une analyse du vocabulaire contenu dans les méthodes de sa zone en se choisissant une liste comme modèle. Les auteurs de listes de fréquence combinée peuvent aussi tirer profit d'un inventaire de ce genre qui leur offre un éventail des compilations déjà existantes. La bibliographie, bien qu'axée sur les travaux allemands, apporte aussi quelques titres sur d'autres langues. Les appendices complètent ces informations surtout pour le chercheur et l'étudiant.

3.3 DEFICIENCES DE L'ANALYSE

3.3.1 Travaux manquants

Notre analyse s'est concentrée sur les ouvrages parus en Amérique du Nord et en Allemagne comme s'il n'y en avait plus ailleurs. Nous n'avons entamé l'analyse que des seuls travaux qui nous ont été accessibles. Nous nous sommes parfois contenté d'un simple compte rendu trouvé dans une revue ou une bibliographie analytique pour certains des travaux manquants qu'on tenait à décrire. Les listes de France (Auber), de Hollande (Schneider) et de Russie nous ont particulièrement manqué.

3.3.2 Grille d'analyse non complétée

Dans l'analyse de certains travaux, plusieurs points prévus dans notre grille (1.4) n'ont pas été considérés. Ceci provient du fait que les auteurs de listes ne donnent pas tous les renseignements. Souvent on ne savait pas d'où étaient choisis les mots d'une liste, puis la grandeur du corpus, l'âge des textes, le critère de sélection manquaient.

3.4 SUGGESTIONS POUR DES ANALYSES ULTERIEURES

3.4.1 Les questionnaires

On pourrait peut-être compléter le présent inventaire en envoyant un questionnaire élaboré à différents centres de recherche ou à des particuliers spécialisés en lexicométrie allemande. Ce procédé s'avère efficace dans certaines recherches comme l'inventaire des langues du monde (C.I.R.B., Université Laval). Mais l'analyse ne doit pas uniquement reposer sur ces questionnaires, il y a lieu de faire des analyses directes à partir des listes elles-mêmes.

3.4.2 Classement

Nous avons classé les listes en deux groupes tel qu'indiqué dans l'introduction (0.2.3). Ce type de classement fait un peu ressortir l'influence que Kaeding a eu sur certains auteurs. Un autre procédé, considérant chaque ouvrage pour lui-même et pour son rapport avec les autres, consisterait à distinguer trois groupes:

- un premier que nous appellerions travaux primaires réunirait et analyserait tous les ouvrages reposant sur des recherches personnelles de l'auteur. C'est le cas du travail de Kaeding et Pfeffer par exemple.

- un second groupe, les travaux secondaires, réunirait et analyserait les travaux fondés uniquement sur les résultats d'autres chercheurs. C'est le cas de la majorité des ouvrages analysés dans la première partie de cette recherche.

- un troisième groupe, les travaux mixtes, réunirait et analyserait les travaux utilisant les deux méthodes ci-dessus. C'est le cas du *Grundeutsch-Mittelstufe* (187) qui repose sur des recherches personnelles et sur des mots tirés d'autres listes.

3.4.3 La valence lexicale et la disponibilité

S'il y a eu des progrès dans les recherches allemandes, tous les critères objectifs de sélection connus à date n'ont pas encore été étudiés pour l'allemand. Aussi une étude sur la valence lexicale serait-elle à entrevoir de façon sérieuse.

Il en est de même de l'étude de la disponibilité que l'on pourrait reprendre en combinant diverses méthodes notamment celle de Bakonyi, dans sa façon de rechercher les mots usuels, celle de Steger et Keil, celle de Pfeffer. Le résultat obtenu pourrait aussi passer sous forme de test de familiarité. Une étude de ce genre permettrait, dans une certaine mesure, de découvrir l'univers familier des enfants. En effet, les listes de mots disponibles et familiers

qu'on sortirait ainsi donneraient les concepts les plus connus et familiers aux enfants.

4. BIBLIOGRAPHIE

4.1 ABREVIATIONS ET SYMBOLES

1 BABEL - *Babel, Journal of Translation.* Melbourne, Australie.

2 BSLP - *Bulletin de la société de linguistique de Paris.* Paris.

3 CDL - *Cahiers de lexicologie.* Paris.

4 CIRB - *Centre international de recherches sur le bilinguisme.* Université Laval, Québec.

5 CJL - *The Canadian Journal of Linguistics / La revue canadienne de linguistique.* Toronto.

6 DNS - *Die Neuren Sprachen.* Francfort.

7 ELA - *Etudes de linguistique appliquée.* Besançon, Paris.

8 GQ - *The German Quarterly.* Cincinnati, U.S.A.

9 HTM - *Helmut Meier.* (131).

10 IRAL - *International Review of Applied Linguistics in Language Teaching.* Heidelberg, Allemagne.

11 IRET - *Institute for Research in English Teaching.*

12 IZFE - *Internationale Zeitschrift für Erziehungswissenschaft.* Berlin.

13 LL - *Language Learning.* Ann Arbor, Michigan, U.S.A.

14 LM - *Les langues modernes.* Paris.

15 LLBA - *Language and Language Behaviour Abstracts.* Ann Arbor, Michigan, U.S.A.

16 MDU - *Monatshefte für deutschen Unterricht.* Madison, Wisconsin, U.S.A.

17 MFDSP - *Monatshefte für deutsche Sprache und Pädagogik.*

18 ML - *Modern Languages.* Londres.

19 MLF - *Modern Language Forum.* Los Angeles, California, U.S.A.

20 MLJ - *Modern Language Journal.* St-Louis, Missouri, U.S.A.

21 MLQ - *Modern Language Quarterly.* Seattle, Washington, U.S.A.

22 MLR - *Modern Language Review.* Cambridge, G.B.

23 MSPE - *Muttersprache.*

24 NZ - *Neuphilologische Zeitschrift.* Berlin.

25 PÄD - *Pädagogik.* Berlin.

26 PBML - *The Prague Bulletin of Mathematical Linguistics.* Prague.

27 PMLA - *Publications of the Modern Language Society of America.* New York.

28 PUF - *Presses Universitaires de France.*

29 RELV - *Revue de l'enseignement des langues vivantes.*

30 RL - *Revue de linguistique.*

31 RLV - *Revue des langues vivantes.* Bruxelles.

32 WUS - *Wörter und Sachen.* Heidelberg.

33 ZFEU - *Zeitschrift für französischen und englischen Unterricht.*

34 ZFNU - *Zeitschrift für neusprachlichen Unterricht.*

35 ZFPH - *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung.* Berlin.

4.2 OUVRAGES

1 AUBER, M. *Guide de fréquence allemand.* Gap: Editions Ophrys. 1953.

2 AUBER, M. *Le vocabulaire pour baccalauréats, selon la fréquence d'emploi des mots.* Gap.

3 *Actes du Sixième Congrès International des Linguistes.* Paris, 1948. P. 83-91.

4 APPELT, E.P. Wortschatz und Kulturkunde, *MDU.* 30(5):264-267, May, 1938.

5 BAGSTER-COLLINS, E.W. Standardization of the Vocabulary in Beginners' Books, *MFDSP.* 18:245-268.

6 BAGSTER-COLLINS, E.W. Observations on Extensive Reading. *GQ.* 3(I):18-27, January 1931.

7 BAKONYI, H. *Die gebräuchlichsten Wörter der deutschen Sprache für den Fremdsprachenunterricht stufenmässig zusammengestellt.* Verlag Ernst Reinhardt, München, 1934. 86 p. (1ère édition Goethe Institut 1933; München 1939).

8 BARON, Samuel. *The Psychological Basis of Vocabulary Teaching in Foreign Language Instruction.* M.A. Thesis, College of the City of New York, 1931. 96 p.

9 BIERWIRTH, H.C. *Words of Frequent Occurrence in Ordinary German.* New York, Holt, 1928. 61 p.

10 BONGERS, H. *The History and Principles of Vocabulary Control.* Wocopi-Woerden, Holland, 1947. P. 360.

11 BRANDT, H.C.C. *German-English Dictionary.* Dans Hauch: German Idiom List. P. 1.

12 BUCHANAN, M.A. et E.D. MACPHEE. *Annotated Bibliography of Modern Language Methodology.* Toronto, University of Toronto Press, 1928. 428 p.

13 BUCHANAN, M.A. *A Graded Spanish Word Book.* Publications of the American and Canadian Committee on Modern Languages, Toronto, 1929. Vol III.

14 *Bulletin d'information du Laboratoire d'analyse Lexicologique.* (Publication du CEVF, Faculté des Lettres et des Sciences Humaines), Besançon, 1960.

15 *Cahiers de Lexicologie.* Paris, 1959.
Vol. 1: Mécanisation et lexicologie, 1959.
Vol. 2: Lexicologie et lexicographie, 1960.
Vol. 3: Colloque International sur la mécanisation des recherches linguistiques, Besançon, 1961.

16 THE CENTRE FOR INFORMATION ON LANGUAGE TEACHING AND THE ENGLISH-TEACHING INFORMATION CENTRE OF THE BRITISH COUNCIL. *A Language-Teaching Bibliography.* Cambridge at the University Press, 1968. 244 p. Vol X.

17 CHATEAUNEUF, Amy. *Study of the Written Vocabulary in a Foreign Language.* Educational Outlook, May 1933. P. 224-234.

18 CLOSSET, Fr. *Didactique des langues vivantes.* Paris, Didier, (sans date). 270 p.

19 COHEN, M. Statistique linguistique. *Actes du Congrès International des Linguistes.* Paris, 1948.

20 COCHRAN, Emory E. Mastering German idioms. *MLJ.* 22(4):274-276, January 1938.

21 COLEMAN, Algernon. *An Analytical Bibliography of Modern Language Teaching, 1927-1932.* The University of Chicago Press, 1933. XIII + 296 p.

22 COLEMAN, Algernon. *An Analytical Bibliography of Modern Language Teaching, 1932-1937.* The University of Chicago Press, 1938. XVIII + 561 p. Vol. II.

23 COLEMAN et FIFE. *An Analytical Bibliography of Modern Language Teaching, 1937-1942.* New York, Columbia University, 1949. XIII + 549 p. Vol. III.

24 CONDON, E.V. Statistics of Vocabulary. Dans *Science.* 1928. 300 p. No 67.

25 *Conférences de l'Institut de Linguistique de l'Université de Paris.* Paris, C. Klincksieck, 1950.

26 HOUSEHOLDER, Fred W. et Sol SAPORTA, ed. *Problems in Lexicography, report.* Conference on Lexicography, 1960 (November 11-12). Mouton, Indiana University, 1962. 286 p.

27 COOPER, R.L. et L. Greenfield, ed. Word Frequency Estimation as a Measure of Degree of Bilingualism. *MLJ.* 53(3):163-166, March 1969.

28 COSERIU, E. Structure lexicale et enseignement du vocabulaire. Dans *Les théories linguistiques et leurs applications*. AIDELA, Conseil de la coopération culturelle du Conseil de l'Europe, Strasbourg, 1967. P. 9-51.

29 DALE, E. et D. REICHERT. *Bibliography of Vocabulary Studies*. Bureau of Educational Research, Ohio State, 1939. 56 p.

30 DIETZE, Hugo. *Methodik des fremdsprachlichen Unterrichts in Handelsschulen*. Leipzig, G.A. Glockner, 1927. P. 148 + 12.

31 DRACH, Erich. *Deutsche Aussprachelehre für den Gebrauch im Ausland*. Diesterweg, Frankfurt/Main, 1931.

32 DROESCHER, W.O. Eine Übersicht über den Stand der deutschen Linguistic. Dans *Deutschunterricht für Ausländer*. Heft ½, München, 1963. P. 27-34.

33 DUBOIS, Jean. Les problèmes du vocabulaire technique. *Lex. 9*: 103-112, 1966.

34 EATON, H.S. *Comparative Word Frequency List*. New York, International Auxiliary Language Association, 1934.

35 EATON, H.S. *Semantic Frequency List for English, French, German, Spanish*. The University of Chicago Press, 1940. 441 p. XXI.

36 EGGELING, H.F. *A dictionary of modern German prose usage*. Oxford, Clarendon Press, 1961. XII + 418 p.

37 EGGERS, Hans. *Beobachtungen zur Häufigkeit deutscher Wortformen*. Dans WW 17, 1967. P. 93-105.

38 EHRKE, Karl. *Methodik des neusprachlichen Unterrichts*. Berlin, F.A. Herbig, 1928. 75 p.

39 EHRKE, Karl. Publications of the American and Canadian Committees on Modern Languages. *DNS* 39(3):227-234, March 1931.

40 ENGEL, E.F. The Use of a Standardized Vocabulary in Beginning German. *MLJ* 15(4):281-291, January 1931.

41 ENGEL, Ulrich. *Sprachwissenschaft und Deutschunterricht*. Dans Deutschunterricht für Ausländer. Heft 5-6, 1968. P. 153-164.

42 ENGELS, L.K. The Fallacy of Word-Counts. Dans *IRAL*. AoÛt 1968. Vol. 6, no 3.

43 ERK, Heinrich. *Ein Rangwörterbuch und viele Fragen.* Dans Deutschunterricht für Ausländer. 1963. P. 86-93, 3/4.

44 FACULTE DES LETTRES ET DES SCIENCES HUMAINES DE STRASBOURG. *Statistique et analyse linguistique.* Paris, P.U.F., 1966. 135 p.

45 FAUCETT, L. *Essential Differences between the Vocabulary of the Interim Report on Vocabulary Selection and Vocabularies prepared for the simplification of English.* Bulletin of the Institute for Research in English Teaching. (127):8-10, October 1936.

46 FERGUSON, Charles et al. *Linguistic Reading Lists for teachers of Modern Languages.* Washington, D.C., 1963. 114 p.

47 FIFE, R.H. *Frequency Word Lists.* MLF, October 1930. P. 125-128, XV.

48 FIFE, R.H. *A Summary of Reports on the Modern Foreign Languages Issued by the Modern Foreign Language Study and the Canadian Committee on Modern Languages.* New York, Macmillan, 1931. VII + 261 p.

49 FINDLAY, J.J. *A Note on the Acquirement of Vocabulary.* ML, Oct. 1927. P. 4-7, IX.

50 FORTOS, J. Word and Idiom Frequency Counts in French and their value. *MLJ.* 15:344-353, Feb. 1931.

51 FRAISSE, P. *Fréquence et familiarité du vocabulaire.* Problèmes de Psycho-linguistique. Paris, P.U.F., 1963. P. 151-167.

52 FRIES, C. et TRAVER. *English Word List, Study of the Adaptability for Instruction Prepared for the Committee on Modern Language.* American Council on Education, Washington, 1940. 92 p.

53 GABKA, Kurt. *Theorien zur Darstellung eines Wortschatzes mit einer Kritik der Wortfeldtheorrie.* Germanistik 9, 1968. Rainer Rath Linguistische Studien: Halle /s, Niemeyer, 1967. 62 p.

54 GAMBLE, W. *Two Lists of Selected Characters Containing All in Bible and Twenty-seven other Books.* Shangai, 1861.

55 GARCIA HOZ, V. *Vocabulario usual, vocabulario común, y vocabulario fundamental.* Consejo Superior de Investiciones Cientificas, Instituto "San José de Calasanz", Madrid, 1953.

56 GELLENTHIEN, Elvira J. *Analysis of First-year German Textbooks.* (Thèse de maîtrise manuscrits), University of Chicago, 1935. 95 p.

57 GIESECKE, E. Cognates, Stem-Meanings and the Vocabulary Problem. *MDU.* 29(2):73-76, Feb. 1937.

58 GLINZ, H. Über Wortinhalte, Wortkörper und Trägerwerte im Sprachunterricht. *IRAL.* 1(63):42-49.

59 GLINZ, H. *Grundbegriffe und Methoden inhaltbezogener Text und Sprachanalyse.* Düsseldorf, pädagogischer Verlag Schwann, 1965. 172 p.

60 GOOSE, André. Quelques réflexions sur les méthodes de la lexicologie. *RLV.* 32:3-13, 1966.

61 GOETHE-INSTITUT, ARBEITSSTELLE FÜR WISSENSCHAFTLICHE DIDAKTIK. *Arbeitsmittel für den Deutschunterricht an Ausländer.* München, 7 Auflage, 1970. 70 p.

62 GOUGENHEIM, G. Statistique linguistique et histoire du vocabulaire. *CDL.* Vol. 2.

63 GOUGENHEIM, G. et al. *L'élaboration du français élémentaire, Etude sur l'établissement d'un vocabulaire et d'une grammaire de base.* Paris, Didier, 1956. XIII + 256 p.

64 GOUGENHEIM, G. *L'élaboration du français fondamental (1er degré).* Nouvelle édition, Paris, Didier, 1964.

65 GRIESBACH, H. und Dora SCHULZ. *Glossar zu Deutsche Sprachlehre für Ausländer.* München, Hueber Verlag, 1958. 79 p.

66 GRIESBACH, H. *1000 Idiomatische Redensarten Deutsch mit Erklärungen und Beispielen.* Verlag Langenscheidt, Berlin, 1969. (6 Auflage), 253 p.

67 GRUTTER, W. Die deutschen Bildwörterbücher im Unterricht. *Deutschunterricht im Ausland, 1937-1938.* (3):1-3, November 1937.

68 GUILBERT, L. De l'utilisation de la statistique en lexicologie appliquée. Dans *ELA.* Paris, Didier, 1963. P. 12..., no 2.

69 GUIRAUD, P. *Les caractères statistiques du vocabulaire: essai de méthodologie.* Paris, P.U.F., 1954 a. 116 p.

70 GUIRAUD, P. *Bibliographie critique de la statistique linguistique.* Utrecht, Spectrum, 1954 b. 123 p.

71 GUIRAUD, P. *Problèmes et méthodes de la statistique linguistique.* Paris, P.U.F., 1960. 145 p.

72 HAASE, A. Wertigkeitsstufen als Grundlagen für eine systematische Wortschatzarbeit und den praktischen Erfolg des Fremdsprachenunterrichts. *DNS.* 1960. 6.

73 HAGBOLT, P. The Relative Importance of Grammar in A German Reading Course. *GQ.* 1:18-21.

74 HAGBOLT, P. *Building the German Vocabulary.* University of Chicago Press, 1928. XVII + 71 p.

75 HAGBOLT, P. Presenting Grammar Inductively. *MLJ.* 12(6):440-445, March 1928.

76 HAGBOLT, P. The Association of the Central West and South Adopts a New Standard Word List. *GQ.* 4(3):118-123, May 1931.

77 HALLIG, Rudolf. *Begriffssystem als Grundlage für die Lexikographie Versuch eines Ordnungsschemas.*

78 HAMMER, Carl. Stress the German-English Cognates! *MLJ.* 41(4): 177-182, April 1957.

79 HAMMOND, Robin T. Pfeffer's Grunddeutsch. *ML.* 47(4):157-160, December 1966.

80 HANDSCOMBE, R.J. The Nuffield Child Language Survey. Dans *Languages and the Young School Child.* Stern H.H., London, Oxford University Press, 1969.

81 HARD, Gerhard. *Zur Mundartgeographie: Ergebnisse, Methoden, Perspektiven.* Beiheft 17 zur Zeitschrift "Wirkendes Wort" Düsseldorf, Pädagogischer Verlag Schwann, 1966. 75 p.

82 HARKIN, Duncan. The History of Word Counts. *BABEL.* 1957. P. 113-124, 3.

83 HARTMANN, R. Bedeutung und Entwicklung der Worthäufigkeitsforschung. *Beiträge zur Sprachkunde und Informations-verarbeitung 4.* Oldenburg Verlag, München, 1964. P. 47-57.

84 HATHAWAY, L.V. A German Error Count - An Experimental Study. *MLJ.* 13:512-533, 1929.

85 HAUCH, E.F. *German Idiom List.* Publication of the American and Canadian Committees on Modern Languages, New York, 1929. XI + 98 p. Vol. X.

86 HEDGCOCK, F.A. Word Frequency in French. *ML.* 17(1):17-18, October 1935.

87 HELBIG, G. et W. SCHENKEL. *Wörterbuch zur Valenz und Distribution deutscher Verben.* VEB Bibliographisches Institut, Leipzig, 1969. 311 p.

88 HELMERS, Hermann. *Didaktik der deutschen Sprache; Einführung in die Theorie der Muttersprache und literarischen Bildung.* Stuttgart, Klett, 1966. 340 p.

89 HENMON, V.A.C. *A French Book Based on a Count of 400,000 Running Words.* Madison (Wisconsin), Bureau of Educational Research, University of Wisconsin (texte ronéotypé), 1924.

90 HENMON, V.A.C. et al. *American Council Alpha German Test Forms A and B.* Yonkers-on Hudson, New York, World Book Co., 1926.

91 HENMON, V.A.C. *Achievement Tests in the Modern Foreign Languages.* New York, Macmillan, 1929. XXVI + 363 p.

92 HENMON, V.A.C. The Vocabulary Problem in the Modern Foreign Languages. *Monatshefte.* 22:33-39, February 1930.

93 HENRION, P. Un mariage malheureux: Statistique et Vocabulaire. *LM.* 43:238-245.

94 HEWITT, T. For Briefer Text - Vocabulary. *GQ.* 9(1):25-26, January 1936.

95 HEWITT, T. A suggested Method for Vocabulary building. *Monatshefte.* 43(4-5):235-236, April-May 1951.

96 HIEBLE, Jacob. Toward a better understanding of the German Vocabulary. *GQ.* 33(1):28-34, January 1960.

97 HOHLFELDT, A.R. Minimum Standard Vocabulary for German. *GQ.* 7(3):87-119, May 1934.

98 HORN, E. *A Basic Writing Vocabulary.* Iowa City, University of Iowa, Monographs in Education, 1926.

99 JADOT, R. *Recherche sur l'importance et la fréquence des mots dans l'étude d'une langue étrangère.* Bruxelles, Archives belges des Sciences de l'Education, 1936. Cahier I, tome I.

100 JONES, L. et J.M. Wepman. *A Spoken Word Count.* Language Research Associates, Chicago, 1966.

101 JUILLAND, A. et CHANG-RODRIGUEZ. *Frequency dictionary of Spanish Words.* The Hague and London, Mouton, 1964. LXXVIII + 500 p.

102 KAEDING, Friedrich Wilhelm. *Häufigkeitswörterbuch der deutschen Sprache.* Steglitz bei Berlin, der Herausgeber, 1897-1898.

103 KAUFHOLZ, E. et al. *Vocabulaire de base de l'allemand.* Paris, 1963.

104 KEIL, Rolf-Dietrich. Einheitliche Methoden in der Lexikometrie. *IRAL.* Heidelberg, 3(2):95-122, May 1965.

105 KELLER, Marie V. The Necessity of Teaching a Basic Vocabulary in Modern Language Work. *MLJ.* 8:35-40, October 1923.

106 KENISTON, Hayward. *A Standard List of Spanish Words and Idioms.* Boston, Heath, 1941. XIV + 108 p.

107 KOEKKOEK, B.J. *J. Alan Pfeffer: Grunddeutsch Basic (spoken) German Word List Mittelstufe.* Zielsprache Deutsch 2-1970. München, p. 99.

108 KÖPKE, Wulf. Viele Fragen an die Sprachstatistik. *Deutsch für Ausländer 14.* 4-6(64):112-124.

109 KROLLMANN, F. et al. Herstellung textbezogener Fachwortlisten mit einem Digitalrechner - ein Verfahren der automatischen Übersetzungs. *Beiträge zur Sprachkunde und Informations-Vorbereitung.* Heft 5, (1965).

110 LADO, R. *Massive Vocabulary Expansion in a Foreign Language Beyond the Basic Course: The Effects of Stimuli, Timing, and Order of Presentation.* Georgetown University, Washington, D.C. 20007, 1967.

111 LAFORGE, Lorne. *La fréquence et ses modalités: répartition et disponibilité.* Communication à la section de linguistique de l'ACFAS, novembre 1965. 15 p.

112 LAFORGE, Lorne. *La Sélection en didactique analytique.* Université Laval, 1970. X + 722 p. (Sous presse).

113 LAMBERT. *German Idioms.* Cité par Hauch, (85, p. 1).

114 LEMMER, Manfred. Deutscher Wortschatz. *Bibliographie zur deutschen Lexicologie.* Halle/s, Niemeyer 1967. 123 p.

115 LEOPOLD, Werner F. Word-Counts and after. *MDU.* 18(8):355-362, December 1936.

116 LEVASSEUR, Karl. Die deutsche Einheitskurzschrift im Lichte der Kaedingschen Häufigkeitsuntersuchungen. Dans *Helmut Meier.* 1926. (131, p. 350).

117 LORGE, I. The English Semantic Count. Cité par Miller, G.A. dans *Langage et Communications*. Paris, P.U.F., 1951. 154 p.

118 LORICHE, R. *Le mot et l'idée*. Editions Ophrys, Librairie Perche, 1965. 130 p.

119 MACKEY, W.F. *Language Teaching Analysis*. London, Longmans, 1965. XI + 554 p.

120 MACKEY et J.G. SAVARD. The Indices of Coverage. *IRAL*. 5(2-3): 71-121, 1967.

121 MACKEY, W.F. Trends and Research in Methods and Materials. Dans *Languages and the Young School Child*. London, Oxford University Press, 1968. P. 69-83.

122 MACKEY, W.F. et al. *Le vocabulaire disponible du français*. Tome I, Didier, 1971 a. 534 p.

123 MACKEY, W.F. et al. *Le vocabulaire disponible du français*. Tome II, Didier 1971 b. 223 p.

124 MARONPOT, R.P. Let's teach and test vocabulary on a one-language basis. *GQ*. 25(1):26-32, January 1952.

125 MAROUZEAU, J. *Lexique de la terminologie linguistique*. Paris, 1943.

126 MARTHALER, Theo. *Das Häufigkeitsprinzip im Sprachunterricht*. Die Neue Schulpraxis 9, 1949. P. 97-103.

127 MATORE, Georges. *La méthode en lexicologie; domaine français*. Paris, Didier, 1953. 126 p.

128 MATTUTAT, Heinrich. *Deutsche Grundsprache, Wort - und Satzlexikon*. Stuttgart, Ernst Klett Verlag, 1963. 488 p.

129 MEIER, Helmut. Dreissig Jahre Zählforschung am deutschen Sprachschatz. *Muttersprache*. 1951. P. 6-14.

130 MEIER, Helmut. Die tausend häufigsten Wortformen der deutschen Sprache. *Muttersprache*. 1952. P. 88-94.

131 MEIER, Helmut. *Deutsche Sprachstatistik*. Vols 1-2, end. rev. and enl. edn. Olms, Hildesheim, 1967. XIII + 422, 150 p.

132 MEIER, Hermann. A Suggested List of 1,000 active German Words. *GQ*. 6:77-90, March 1933.

133 MEIER, Hermann. *The 1,000 most frequent German Words*. New York, Oxford, 2nd ed., 1933. Reprinted 1947. 24 p.

134 MELDAU, Rudolf. *Kleines Deutschlandbuch für Ausländer. Wichtige Sachgebiete und ihr Wortschatz.* Munich, Max Hueber Verlag, 2nd ed., 1959. 77 p.

135 MELZER, Friso. Das Wort in den Wörten. Die deutsche Sprache im Dienste der Christus-Nachfolge. *Ein theophilologisches Wörterbuch Tübingen.* Mohr <Siebeck>, 1965. XVI + 372 p.

136 MENZERATH, P. Die Bedeutung der sprachlichen Geläufigkeit oder der formalen sprachlichen Beziehung für Reproduction. *Zeitschrift für Psychologie 48.* 1908.

137 MENZERATH, P. *Die Architektonik des deutschen Wortschatzes.* [phonetische Studien, 3]. Bonn/Hannover/Stuttgart, Ferd. Dümmlers Verlag, 1954. VIII + 131 p.

138 MERAS, A.A. et M. MILLER. *Ein Wortschatz.* Boston, Heath, 1916. 32 p.

139 MEYER, L. (Alld gotique) Décompte du mot parlé, Berlin, 1869. Dans *Hartmann.* (83).

140 MICHEA, René. La normalisation des vocabulaires scolaires. *RLV.* 5:283-299, 1948.

141 MICHEA, René. Le vocabulaire de base aux examens. *LM.* 43:135-140, 1969.

142 MICHEA, René. Introduction pratique à une statistique du langage. *LM.* 43:173-186; 314-322, 1949.

143 MICHEA, René. Vocabulaire et culture. *LM.* 44:187-192, 1950.

144 MICHEA, René. La culture par la langue. *LM.* 44:328-335, 1950.

145 MICHEA, René. Rapports de la fréquence avec la forme, le sens et la fonction des mots. *LM.* 1951. P. 191-196.

146 MICHEA, René. Vocabulaire et physiologie. *LM.* Mai-juillet 1952. P. 227-232.

147 MICHEA, René. Le vocabulaire de base en France et à l'étranger. *LM.* 1952. P. 395-397.

148 MICHEA, René. Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *LM.* 47:338-344, 1953.

149 MICHEA, René. Limitation et sélection du vocabulaire dans l'enseignement actif des langues vivantes. *RLV.* 22, 5(56): 467-474.

150 MICHEA, René. Le vocabulaire allemand d'utilité scolaire et générale. *RLV.* 22, 6(56):560-573.

151 MICHEA, René. Traduction scientifique et organisation du vocabulaire. *LM.* 1959. P. 486-495.

152 MICHEA, René. *Vocabulaire allemand progressif.* Didier, Paris, 1959. 304 p.

153 MICHEA, René. Les vocabulaires fondamentaux. Dans *Recherches et techniques nouvelles au service de l'enseignement des langues vivantes.* Conseil de la Coopération culturelle du Conseil de l'Europe, Strasbourg, 4(3):21-36, 1964.

154 MILLER, G.A. A Statistical Description of Verbal Learning. *Psychometrika.* (17):309-396, 1952.

155 MINISTERE DE L'EDUCATION NATIONALE. *Le français élémentaire.* Paris, Publication du Centre national de documentation pédagogique, 1954.

156 MOREAU, René. Initiation à la méthode statistique en linguistique. *Bulletin d'information du Laboratoire d'analyse lexicologique.* Besançon, Publication du CEVF, 6:1-28, 1962.

157 MORGAN, B.Q. *German Frequency Word Book based on Kaeding's* Häufigkeitswörterbuch der deutschen Sprache. New York, The Macmillan Company, 1928. XV + 87 p.

158 MORGAN, B.Q. A minimum Standard Vocabulary for German. *MLJ.* 18(3):145-152, December 1933.

159 MORGAN, B.Q. Some Frequency Counts in German Grammar. *Monatshefte.* 32:117-119, 1940.

160 MORGAN, B.Q. et L.M. OBERDECK. *Active and Passive Vocabulary.* Publications of the American and Canadian Committees on Modern Languages, 1930. P. 211-221, vol. XVII.

161 MORGENTHALER, R. *Statistik des Neutestamentlichen Wortschatzes.* Zürich-Ffm, 1958.

162 MOSER, Hugo. Entwicklungstendenzen des heutigen Deutsch. *Der Deutschunterricht.* 1954. P. 87-107, VI.

163 MOSER, Hugo. *Neuere und neueste Zeit: Von den 80er Jahren des 19. Jahrhunderts zur Gegenwart. Deutsche Wortgeschichte.* Ed. by F. Maurer and F. Stroh., vol. II, p. 445-560, ended, Berlin: de Gruyter, 1959.

164 MOSSE, F. Détermination et progression du vocabulaire dans l'enseignement des langues. *LM.* Paris, 1948.

165 MUEGGE, M.A. *German Word-frequency vocabulary.* Pitman, 1946.

166 MULLER, Charles. Fréquence, dispersion et usage: à propos des dictionnaires de fréquence. *CDL.* 6(2):33-42, 1965.

167 MULLER, C. *Initiation à la statistique linguistique.* Paris, Larousse, 1968. 248 p.

168 MULLER, Siegfried H. *The World's living languages.* Frederick Ungar Publishing Co., 1964. XII + 212 p.

169 MULLER, G. Elementarstufe und Umgangssprache. *DNS.* 1964. 8.

170 NAIDU, B.K. *Kannada (Canarese) Word List.* State Adult Education Council, Mysore, India, 1947.

171 NABHOLZ, J. Teaching German Vocabulary through Etymology. *MLJ.* 37(7):347-350, November 1953.

171a NASH, Rose. *Multilingual Lexicon of Linguistics and Philology: English, Russian, German, French.* University of Miami Press, 1968. 390 p., XXVI.

172 *New York State Syllabus.* Tentative Syllabus in Modern Foreign Languages, Albany, University of the State of New York, 1930.

173 NEVENSCHWANDER, Anna. Objectives of a Second Year German Course. *MLJ.* 22(3):209-212, December 1937.

174 NOBLE, C.E. The Meaning-Familiarity Relationship. *Psychological Review.* 60:89-98.

175 NOSTRAND, H.L. et al. *Research on Language Teaching. An Annotated international Bibliography 1945-64.* University of Washington Press, 1965. XXI + 373 p.

176 OGDEN, C.K. *The Basic Words: A Detailed Account of their Uses.* Londres, Psyche Miniatures, 1933.

177 OEHLER, Heinz. *Grundwortschatz Deutsch: Essential German: Allemand fondamental.* Stuttgart, Klett, 1966. 235 p.

178 OEHLER, Heinz. Grundwortschatz Deutsch. *IRAL.* 5:216-219, 1967.

179 OEHLER, H. et al. *Grundwortschatz in 6 Sprachen: Deutsch-Englisch-Französisch-Italienisch-Spanisch-Russisch.* Stuttgart, 1969.

180 ORTMANN, A.A. A Study in First-Year German Vocabulary. *GQ.* 8(3):119-128, May 1935.

181 PALMER, H.E. Interim Report on Vocabulary Selection, Tokyo, IRET, 1930.

182 PALMER, H.E. The Testing of the Word-Lists. *Bulletin of the Institute for Research in English Teaching.* (85):1-2, June 1932.

183 PALMER, H.E. and M. WEST. Discussion of Word Frequency. *ML.* 18(3):136-138, March 1937.

184 PFEFFER, J. Alan. *Dictionary of Everyday Usage English-German, German-English.* New York, 1945. XXVI + 369 p. 504.

185 PFEFFER, J. Alan. Grunddeutsch: Werden und Wesen. *GQ.* 35: 179-186, März 1962.

186 PFEFFER, J. Alan. *Evolution of the Basic (Spoken) German Word List, Level I.* (Preliminary edition, multigraphed, University of Pittsburgh, 1963.

187 PFEFFER, J. Alan. *Basic (Spoken) German Word List Grundstufe.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964. 79 p. IX.

188 PFEFFER, J. Alan. *Index of English Equivalents for the Basic (Spoken) German Word List, Grundstufe.* Prentice-Hall, Inc. 1965. VII + 107 p.

189 PFEFFER, J. Alan. *Basic (Spoken) German Idiom List, Level I.* Prentice-Hall, 1967. IX + 91 p.

190 PFEFFER, J. Alan. Grunddeutsch. *Zeitschrift für deutsche Sprache.* 1969. P. 132-141.

191 PFEFFER, J. Alan. *Grunddeutsch Basic (Spoken) German Word List, Mittelstufe.* Preliminary Edition, Pittsburgh, Institute for Basic German, 1970. 112 p.

192 PFEFFER, J. Alan. *Basic (Spoken) German Dictionary for Everyday Usage.* Prentice-Hall, 1970. XV + 350 p.

193 PFEFFER, J. Alan. La mécanisation dans la préparation statistique de l'allemand fondamental. *SDL.* P. 90-93.

194 PLATH, W. Mathematical Linguistics. *Trends in European and American Linguistics.* 1930-1960, Utrecht 1961.

195 PREHN, A.H. *Practical Guide to German Vocabulary.* Oxford University Press, 1912. 276 p.

196 PRICE, W.R. What Price Vocabulary Frequencies? *GQ*. 2:1-5, January 1929.

197 PURIN, C.M. *A Standard German Vocabulary of 2,000 Words and Idioms Selected from frequency Counts and Illustrated in typical sentences*. Chicago, University of Chicago Press, 1931. XVI + 195 p.

198 PURIN, C.M. *A Standard German Vocabulary of 2,932 Words and 1,500 Idioms*. Boston, Heath, 1937. VI + 185 p.

199 REIFLER, E. MT linguistics and MT lexicography at the University of Washington. *High Frequency Concepts, German Substantive Compounds and MT*. Seattle, 1960.

200 REY, A. *La lexicologie*. Paris, Klincksieck, 1970. 323 p.

201 RICHARDS, J.C. A Psycholinguistic Measure of Vocabulary Selection. *IRAL*. VIII/2, May 1970.

202 RICHARDS, J.C. *Word Familiarity as an Index of Vocabulary Selection, with Indices for 4495 English Nouns*. Québec, Les Presses de l'Université Laval, (1971). 375 p. (Sous-presse).

203 ROBINSON, Janet O. *An Annotated Bibliography of Modern Language Teaching*. Books and Articles 1946-47, London, Oxford University Press, 1969.

204 RODRIGUEZ, Bov and Juana A. MENDEZ. *Recuento de vocabulario español*. Puerto Rico, Universidad de Puerto Rico, 1952. Vols I and II.

205 ROGER, Georges. Pour un enseignement plus méthodique et plus scientifique des langues modernes. *LM*. 29(3):172-182, April 1931.

206 ROSE, E. Vocabularies in German Textbooks. *MLJ*. 17(5):334-341, February 1933.

207 ROSENZWEIG, M.A. Comparison among Word-Association Responses in English, French, German and Italian. *American Journal of Psychology* 74. 1961.

208 ROULSTON, R.B. A Plea for Mor and Better Special Vocabularies. *MLJ*. 13(4):304-307, January 1929.

209 RYDER, Frank G. A Suggestion for Elementary German. *LL*. 3(1-2):3-13, January-June 1950.

210 SAVARD, J.G. *La valence lexicale*. Paris, Didier, 1970. 236 p.

211 SAVARD, J.G. et J.C. RICHARDS. *Les indices d'utilité du vocabulaire fondamental français.* Les Presses de l'Université Laval, 1970. V + 159 p.

212 SCHANZE, Helmut. *Index zu Novalis Heinrich von Ofterdingen.* Frankfurt/Main, Athenäum Verlag, 1968. 193 p.

213 SCHANZE, Helmut. Index zu Heinrich von Kleist. *Sämtliche Erzählungen, Erzählvarianten, Anekdoten.* Frankfurt/Main, Athenäum Verlag, 1969. 330 p.

214 SCHERER, G. A.C. *Word Frequency in the Modern German Short Story.* University of Colorado, Boulder 80304, 1965.

215 SCHINNERER, O. et H.G. WENDT. A Suggested List of 1000 Active German Words. *GQ.* 6(2):77-90, March 1933.

216 SCHMID, Otto. *Deutsch - falsch und richtig; Beispiele aus der heutigen Alltagssprache.* Munich, Max Hueber, 1963. 56 p.

217 SCHIMDT, Peter. *Index Beiheft Der Wortschatz von Goethes <<Iphigenie>> Analyse der Werk - und Personensprache mit EDV-Hilfe. Mit Wortindex, Häufigkeitswörterbuch und Wortgruppentabellen.* Frankfurt/Main Athenäum Verlag, 1970. 149 p.

218 SCHMOLL, G. Leipzig Verlag für Buch - und Bibliothekwesen, 1967. 160 p.

219 SCHNEIDER, J. *Deutscher Wortschatz für niederlandische Schüler.* Batavia, Noordhoff-Kolff, 1937.

220 SCHNELLE, H. Über den Stand der Forschung zur automatischen Sprachbearbeitung im deutschen Sprachraum. *Beiträge zur Sprachkunde und Informationsverarbeitung.* Heft 2, München und Wien, 1963.

221 SCHULTZE, W. *Der Wortschatz in der Grundschule.* Stuttgart, 1956.

222 SCHWERTE, H. et H. SCHANZE. *Indices zur Deutschen Literatur.* Germanistisches Institut der Rhein-Westf. Technischen Hochschule Aaches, Lehrstuhl für neuere deutsche Literaturgeschichte, in Zusammenarbeit mit dem Rechenzentrum der Rhein-Westf. Technischen Hochschule, Frankfurt/Main Athenäum.

223 SDUN, Winfried. Zum Wortschatz der Grundstufe I. *Deutsch für Ausländer.* 15(1-2):37-45, 1965.

224 SILIAKUS, H.J. The 1,000 most useful words in German: The Adelaide List and its compilation. *BABEL.* (27):19-21, October 1964.

225 SILIAKUS, H.J. et Werner SCHLICK. Untersuchungen zur deutschen Worthäufigkeit. *Deutsch für Ausländer.* 15(1-2):46-51, 1965.

226 SIMMONS, L.V.T. A Vocabulary Count Based on Three German Dramas. *MLJ.* 14(1):33-36, October 1929.

227 SLABY, Helmut. *Deutscher Wortschatz in Sachgebieten.* Frankfurt /Main, Moritz Diesterweg, 1965. 95 p.

228 SPALDING, Keith. *An Historical Dictionary of German figurative usage.* Oxford, Blackwell, 1952.

229 SPANG, HANSSEN, Henning. Fini et infini dans le vocabulaire. *Langages* 6. 1967. P. 100-105.

230 SPITZ BARDT, Harry. Zur Entwicklung der Sprachstatistik in der Sowjetunion. *WZUJ* 16. 1967. P. 571-590.

231 *The Statesman's Year Book 1970/1971.* Edited by John Paxton Macmillan, St. Martin's Press, London, WC2R LF, 107th edition 1970. XXVI + 1557 p.

232 STEGER, Hugo. Über Dokumentation und Analyse gesprochener Sprache. *Zielsprache Deutsch.* Heft 1, 1970. P. 13.

233 STEGER, H. et Maria KEIL. *Wortliste für das VHS - Zertifikat Deutsch als Fremdsprache.* Frankfurt/Main, 1971. IX + 132 p.

234 STEINER, O. *Hochdeutsch und Mundart bei Einheimischen und Neubürgern der Kreise Bamberg und Northeim im Jahre 1954 Ergebnisse einer Schulkindererhebung.* Phon 1957. P. 146-156, cité par Zwirner, Eberhard 1962. P. 46 (278).

235 STEINFELDT, E. *Russian word count: 2,500 words most commonly used in Modern Literary Russian: a guide for teachers of Russian.* Moscow, Progress, 1965. 228 p.

236 SWENSON, Rodney. *A Frequency Count of Contemporary German Vocabulary Based on Three Current Leading Newspapers.* University of Minnesota, Minneapolis, Minn., 55455, 1967.

237 TAYLOR, R. et W. GOTTSCHALK. *A German-English dictionary of Idioms.* Munich, Max Hueber, 1960. 597 p.

238 THIESSEN, J. Studien zum Wortschatz der Kanadischen Mennoniten. Thesis, University of Marburg, 1963. Dans *Revue Canadienne de Linguistique*. 15(1):59, 1969.

239 THORNDIKE, E.L. *A Teacher's Word Book of the 20,000 Words Found most Frequently and Widely in General Reading for Children and Young People*. New York, Teacher's College, Columbia University, 1932.

240 THURNAU, H.C. *Vocabulary-Building German for Beginners*. New York, Crofts, 1927. XXVI + 310 p.

241 TOLLENAERE, F. de. Die Automatisierung in der Lexikologie. Dans *Beiträge zur Sprachkunde und Informationsverarbeitung*. Edit. Par Herbert Marchl, Munich, R. Oldenbourg, 1963. P. 33-44, t. I.

242 TOLLENAERE, F. de. *Nieuwe Wegen in de Lexikologie, Amsterdam*. Noord Hollandsche Uitgeversmaatschappij, 1963.

243 TOLLENAERE, F. de. Un dictionnaire historique de la langue allemande: trésor des frères Grimm. *CDL*. 6(1):110, 1965.

244 TRNKA, B. *A Tentative Bibliography*. Publication of the UNESCO Committee on Linguistic Statistics, Brüssel, 1950.

245 TWADDELL, W.F. [Décompte du mot parlé (groupe de consonnes)], Kopenhagen 1938. Cité par Hartmann (83).

246 TWADDELL, W.F. *Language* XXX. 1954 . P. 593-597.

247 UHLIG, W., CHATELANAT, Ch. et J.B. LANG. *Vocabulaire de base allemand-français*. Collection Wir Sprechen Deutsch, Lausanne, Payot, 1963. 128 p.

248 VAIL, C.D. Basic Vocabulary Studies. *GQ*. 5(3):123-130, May 1932.

249 VAIL, C.D. *Basic Word and Idiom Lists*. Albany, University of the State of New York Press, 1933. 22 p.

250 VAIL, C.D. Frequency Word Lists. *GQ* 8. 1935.

251 VAIL, C.D. A Word List Correlation. *GQ*. IX(1):10-16, January 1936.

252 VERLEE, L. Quelques considérations sur le vocabulaire de base. *RLV*. XX, 3.

253 VINETTE, R. *Echelle de vocabulaire et d'orthographe, partie du maître*. Montréal, Centre de psychologie et de la Pédagogie, 1949. 119 p.

254 WADEPUHL, Walter. German Idioms. *GQ*. I(2):68-73, February 1928.

255 WADEPUHL, W. et B.G. MORGAN. *Minimum Standard German Vocabulary*. New York, Appleton-Century-Crofts, 1934. VII + 90 p.

256 WÄNGLER, H.H. *Rangwörterbuch hichdeutscher Umgangsprache*. Marburg, Elwert Verlag, 1963. 67 p.

257 WASSERZIEHER, Ernst. *Woher? Ableitendes Wörterbuch der deutschen Sprache*. 14th ed. by Werner Betz, Bonn, Dümmler, 1959. 441 p.

258 WEIJNEN, A. *Laut und Frequenz der Wörter*. Xeme Congrès International des Linguistes, Bucarest, 28 août - 2 septembre 1967. 402 p.

259 WEISSENBORN, J. Statistique sur la structure de la phrase allemande. *Etudes de Linguistique Appliquée*. Vol. 3, 1964. Dans 8°, p. 163.

260 WENGRAF, William L. *The Applicability of German Songs in Teaching German in Junior and Senior High Schools*. Unpublished Master's Thesis, College of the City of New York, 1933. 111 p.

261 WEST, Michael. Speaking Vocabulary in a Foreign Language (1,000 words). *MLJ*. 14(7):509-521, April 1930.

262 WEST, Michael. The present position in Vocabulary Selection. *MLJ*. 20(6):433-437, March 1937.

263 WEST, Michael. Vocabulary Selection. Dans *The Yearbook of Education*. London, University of London Institute on Education, 1940.

264 WEST, Michael. Vocabulary Selection and the minimum adequate Vocabulary. *English Language Teaching*. VIII, 4.

265 WEST, Michael. *A General Service List of English Words*. London, 1953.

266 WHEELOCK. *List of German Words*. State Department of Education, University of the State of New York.

267 WHEELOCK. *German Word List for Heath's Wall Chasts, with reduced Fac-similes of the charts*. Heath, 23 p.

268 WILLEY, Norman. German Idioms. *MLJ*. 24(3):221-226, December 1939.

269 WOHLGEMUTH-BERGLUND, Gisela. *Wort für Wort*. München, Max Hueber Verlag, 1970. 332 p.

270 WOLF, Friederich. *10,000 Wörter - A German Vocabulary for Students.* München, Max Hueber Verlag, 1969. 140 p.

271 WOOLEY, E.O. *Some Experiments in Teaching German Vocabulary.* Bull. Indiana University. 13(4):21-27, September 1937.

272 ZELLWEGER, Rudolf. *Le vocabulaire du bachelier 3,000 mots allemands choisis et présentés.* Payot, Lausanne, 1961. 80 p.

273 ZEYDEL, Edwin H. On Special Vocabularies. *Monatshefte für deutschen Unterricht.* 20(1):22-26, January 1928.

274 ZIMMERMANN, Jon. Word Frequency in the Modern German shorter narrative. Unpub. diss. University of Colorado, 1967. *GQ.* 44(2):268, March 1971.

275 ZIPF, G.K. Selected Studies of the principle of Relative frequency. Dans *Language.* Cambridge, Mass., 1932.

276 ZIPF, G.K. *The Psycho-Biology of Language.* 1935.

277 ZIPF, G.K. *Human behaviour and the principle of least effort. An Introduction to Hyman Ecology.* Cambridge 1949. XI + 573 p.

278 ZWIRNER, Eberhard. *Deutsches Spracharchiv 1932-1962 Geschichte, Aufgaben und Gliederung, Bibliographie.* Münster Westfalen, Aschendorffsche Buchdruckerei, 1962. 66 p.

5. APPENDICES

5.1 APPENDICE A

Tableau des langues traitées ou en cours de traitement lexicométrique

5.2 APPENDICE B

Principaux centres de traitement

5.3 APPENDICE C

Petit lexique de la lexicométrie

5.4 APPENDICE D

Calculs et notes sur le critère de l'usage

5.5 APPENDICE E

Index des noms d'auteurs parus dans le texte

5.6 APPENDICE F

Résumé du travail

5.1 APPENDICE A

L'appendice A relève les noms des langues du monde ayant déjà fait l'objet d'un traitement lexicométrique.

Les renseignements proviennent de différentes sources:

1 - R.D. Keil, *Einheitliche Methoden der Lexikometrie*

2 - R. Hartmann (83)

3 - C. Muller (166)

4 - S. Muller (168) ± L. Massignon, *Annuaire du monde musulman*, PUF, 1954

5 - La Revue Canadienne de Linguistique 15-1, 1969, p. 50

6 - Personnelles

Des trente et une langues, neuf seulement ne sont pas des langues Indo-européennes.

5.2 APPENDICE B

L'appendice B est loin d'être exhaustif. Il ne donne que quelques noms de centres ou ceux des particuliers auprès de qui on pourrait trouver les premiers renseignements au sujet de la lexicométrie. Nous avons divisé l'appendice en trois sections.

1 - Les Centres spécialisés en d'autres langues que l'Allemand;

2 - Les Centres spécialisés en allemand;

3 - Les noms des particuliers.

5.3 APPENDICE C

L'appendice C présente un certain nombre de termes que nous avons le plus rencontrés au cours de nos lectures. Nous espérons que cette liste succincte sera de quelqu'utilité à l'étudiant non-bilingue qui risque de tomber sur l'un ou l'autre des termes en lisant quelque revue ou article sur la lexicométrie.

Dans *Language Teaching Analysis* (119, 30-32), le professeur Mackey montre l'importance d'une étude de la terminologie linguistique. Afin d'éviter les dangers d'isolation, la branche naissante de la lexicométrie devrait peut-être essayer de se faire un lexique...

Chaque lexicométriste devrait, en créant ou utilisant un nouveau terme, en donner le sens et/ou le correspondant le plus proche en d'autres langues.

La plupart des termes que nous rapportons ici viennent de: W.O. Droescher (no 32), Heinrich Erk (no 43), B.J. Koekkoek (no 107), *IRAL*. V/2-3, July 1967 et Lado, Robert, *Moderner Sprachunterricht, eine Einführung auf Wissenschaftlicher Grundlage*, München, Hueber, 1967, 325 p. et de Michea, René.

f = féminin
m = masculin
n = neutre

APPENDICE A

	Langue	Pays d'origine de la langue	Nombre de locuteurs par millions	Auteurs et/ou chercheurs	Année	Importance du corpus
1	Allemand	Allgne, Autriche, Suisse	100	Kaeding	1898	10,910,777
				Pfeffer	1962-	600,000
2	Anglais	Roy.-Uni, U.S.A.	265	Horn	1922	865,000
3	Arabe (Egyptien)	Egypte	91	Landau	1959	
4	Basaa (MBcMc)	Cameroun	0.5	Njock	1969-	150,000
5	Berbere	Mauritanie	0.013	Destaing	1920	
6	Bengali	Pakistan, Inde	515	Chaudhury	1931	
7	Chinois	Chine	515	Gamble	-1850	
8	Danois	Danemark	5	Mogensen	1957	
9	Espagnol	Cuba, Espagne	138	Buchanan	1927	1,200,000
10	Français	France, Canada	65	Chey@leur	1929	1,183,000
11	Grec	Grèce	7	Comenius	1665	
12	Hindi	Inde, Pakistan	185			
13	Hongrois	Hongrie	13	Univ. Debrecen	1965	
14	Irlandais	Irlande	0.1	Huallachâin	1964	600,000
15	Islandais	Islande, Iles Feroe	0.16	Sigurdson	1940	
16	Italien	Italie	55			
17	Kannada	Inde (Mysore)	18	Naidu	1947	
18	Latin			Masterman	1947	
19	Malayo-Polynesiennes	Malaysia-Polynesia		Dyen	1962	
20	Mexicain (Mahuat)	Mexique	0.5	Prywes, Gray	1962	
21	Neerlandais	Hollande	17	De la Court	1937	1,000,000
22	Norvegien	Norvège	4	Haugen	1942	3,142,000
23	Persan	Afghanistan, Iran	20	Kramsky	1939	
24	Portugais	Brésil, Portugal	85	Brown et al.	1945	1,200,00
25	Roumain	Roumanie, URSS	20	Juilland	1965	
26	Russe	U.R.S.S.	135	Josselson	1953	1,000,000
27	Suédois	Suède, U.S.A.	9	Hassler-Göransson	1930-	
28	Tchèque	Tchécoslovaquie	10			
29	Thai	Thailand, Chine	24			
30	Tokelau	Ile Tokelau	0.002	Boardman	1969	
31	Urdu	Pakistan, India	100	Barker et al.	1969	

APPENDICE B

I

1. THE ACADEMY OF THE HEBREW LANGUAGE (Jerusalem)

En 1961, elle projetait un dictionnaire historique de la langue nationale. Elle a travaillé quelques index et concordances d'un texte Talmudie.

2. CENTRE D'ETUDE DU VOCABULAIRE FRANCAIS [CEVF] à Besançon (France)

En 1957, installation à l'Université de Besançon.

But principal: Sortir un inventaire général et complet du vocabulaire français.

Mais il fait aussi des travaux pour d'autres langues (entre autres l'indexation et l'étude statistique de la recherche de Pfeffer sur le Grunddeutsch ont été réalisés dans ce centre).

Publications: Bulletin d'information du laboratoire d'analyse lexicologique.

3. CENTRO PER L'AUTOMAZIONE DELL'ANALISI LETTERARIA [CAAL] à Gallarate (Italie)

Le Centre est ouvert depuis le 17 décembre 1956.

Travaux principaux: dépouillement Opera Ommia de Thomas d'Aquin. En plus des index, on prépare aussi des statistiques sur la fréquence des graphèmes et de leurs combinaisons.

4. CENTRE INTERNATIONAL DE RECHERCHES SUR LE BILINGUISME [CIRB] Université Laval (Québec - Canada)

Fondé en 1967.

Les principaux travaux de lexicométrie sont produits par les professeurs W.F. Mackey, J.G. Savard, J.C. Richards, L. Laforge.

II

1. ARBEITSSTELLE FÜR WISSENSCHAFTLICHE DIDAKTIK

Ce centre a été créé en 1965 à l'Institut Goethe à Munich (Allemagne).

Rôle: Il veut suivre les progrès des recherches en linguistique générale et appliquée et particulièrement dans les recherches sur l'allemand contemporain. Il prépare des matériaux utilisables dans l'enseignement de l'allemand à l'étranger. Des recherches lexicométriques sur l'allemand scientifique y sont en cours.

2. DEUTSCHES SPRACHARCHIV

Fondation: 1932.

Buts: Enregistrement et évaluation de la langue parlée (travaux faits sur bandes magnétiques, films sonores, disques, etc...). Depuis 1962, les résultats des travaux se rapportant au lexique, à la grammaire et à la syntaxe de la langue allemande sont publiés dans la revue *Idiomatica* chez l'éditeur Niemeyer, à Tübingen, Allemagne.

3. DER DEUTSCHE SPRACHATLAS ou DAS FORSCHUNGSINSTITUT FÜR DEUTSCHE SPRACHE an der UNIVERSITÄT MARBURG (LAHN)

Fondation: 1876 par Georg Wenker.

But principal: Etude de la langue parlée par les dialectes. Pour la collection et la conservation du matériel de recherche, l'Institut utilise deux méthodes: les questionnaires et les interviews. Il dispose déjà d'un atlas des sons et des formes des parlers allemands.

4. GOETHE ARCHIV (GOETHE-WÖRTERBUCH)

Le travail commence après la deuxième guerre.

But: promouvoir l'enseignement de la littérature et de la compréhension de l'oeuvre de Goethe au point de vue sémantique. Le travail se fait à Hamburg, Berlin et Tübingen. Ils font aussi des études de fréquence dans l'oeuvre de Goethe.

5. INSTITUT FÜR DEUTSCHE SPRACHE

Fondation: 29 avril 1964 à Mannheim (Allemagne).

But: Une des tâches les plus importantes était de faire une collection de textes qui serait le fondement de toute recherche scientifique dans le domaine linguistique. Le corpus comprend des textes des années 1945 et après. Le centre de Mannheim s'occupe surtout de la langue écrite et dispose d'un corpus d'environ 1,6 millions de mots courants. Le Centre de Freiburg travaille sur la langue parlée et son corpus dépasse 600 000 mots. Le Centre de Bonn qui travaille sur les textes des journaux a aussi compilé plus de 1,2 millions de mots.

6. INSTITUTE FOR BASIC GERMAN (INSTITUT FÜR GRUNDDEUTSCH)

Fondation: 1960 à l'Université de Buffalo (New York, U.S.A.) avec l'aide du gouvernement américain.

Depuis 1962, il est à Pittsburg (Pennsylvanie). Le fondateur et directeur est le professeur J.A. Pfeffer.

Buts:

a) Immédiats: Analyse de la langue contemporaine d'après les critères de la fréquence et de la "Geläufigkeit";

b) A long terme: Promouvoir l'enseignement de la langue parlée et écrite contemporaine.

De 1962 à 1970, l'Institut a produit huit ouvrages et importants articles. (Voir Bibliographie 184-193) sur la lexicométrie allemande.

7. LEHRGEBIET DEUTSCH FÜR AUSLÄNDER

RWTH 51 Aachen, Templergraben (Allemagne)

8. ZENTRALSTELLE FÜR MASCHINELLE DOKUMENTATION (ZMD)

Frankfurt / Niederrad, Harriotstrasse (Allemagne)

Travaille sur les vocabulaires spécialisés (Fachsprachenwortschatz)

III

1. JUILLAND, A., de Stanford University, prépare des dictionnaires de fréquence de mots, de morphèmes, de syllabes des cinq principales langues romanes.

2. TOLLENAERE, F. de, fait des travaux en lexicométrie depuis 1961. Dans son livre (Voir Bibliographie, 242, p. 136 ff.). Il cite de nombreux centres spécialisés dans ce genre de travail.

3. WISBEY, R. (Lecturer in German, Cambridge University) travaille avec l'Institut des mathématiques de l'Université sur un "Computer" EDSAC 2 des textes allemands du moyen âge (surtout du onzième et douzième siècle).

4. KEIL, Rolf Dietrich,
Sprachenschule der Bundeswehr
535 Euskirchen

C'est Keil qui a lancé le terme "Lexikometrie".

APPENDICE C

	Allemand		Français		Anglais
1	Anwendbarkeit	(f)	disponibilité	(f)	availability
2	Ausdehnung	(f)	extension	(f)	
3	Begriffswort	(n)	mot thématique	(m)	content word
4	Dingwort	[3](n)			
5	Eingang ≠ Ausgang	(m)	encodage ≠ décodage	(m)	input ≠ output
6	Einschliessung	(f)	inclusion	(f)	inclusion
7	Ersetzbarkeit	(f)	substitution	(f)	substitution
8	Erstsprache	(f)	langue première	(f)	first language
9	Erweiterung	(f)	expansion	(f)	
10	Frequenz	(f)	fréquence	(f)	frequency
11	Formwort	(n)	mot athématique	(m)	structure word
12	Glottisverschluss	(m)	coup de glotte	(m)	glottal stop
13	Grammatische Funktionswörter		mots outils		function/grammatical words
14	Grunddeutsch	(n)	allemand fondamental		basic german
15	Grundwortschatz	(m)	vocabulaire fondamental		basic vocabulary
16	Häufigkeit	[10](f)			
17	Häufigkeitszählung	(f)	étude de fréquence		frequency count

	Allemand		Français		Anglais
18	Kriterium	(n)	critère	(m)	criterion
19	Lexikalische Einheit	(f)	unité lexicale	(f)	
20	Lexikometrie	(f)	lexicométrie	(f)	lexicometrics
21	Lochkarte	(f)	carte perforée	(f)	punched card
22	Mehrwertigkeit	(f)	valence	(f)	coverage
23	Redewendung	(f)	tournure idiomatique	(f)	idiom
24	Sprachformen		formes de langage		sentence patterns [Michea 1950]
25	Sprachstatistiche Erhebung	(f)			statistical linguistic survey
26	Stammvokabular		vocabulaire de base	(m)	
27	Stichwort	(n)	mot clef	(m)	keyword
28	Strukturwort	(n)	mot outil	(m)	function word
29	Themenkreis	(m)	centre d'intérêts		topics
30	Valenz der Verben cf. 1.5.2.5				
31	Valence lexicale cf. 1.5.2.5				
32	Verfügbares Wort	(n)	mot disponible	(m)	available/utility word
33	Verfügbarkeit	[1](f)			
34	Verteilung	(f)	répartition/ distribution	(f)	range/distribution
35	Vertrautheit	(f)	familiarité	(f)	familiarity
36	Vertrautheitsgrad	(m)	degré (m) de familiarité		

	Allemand		Français		Anglais
37	Wirkwörter		mots grammaticaux		operators
38	Wortschatz	(m)	vocabulaire	(m)	vocabulary
39	Wortschatzauswahl	(f)	selection (f) du vocabulaire		
40	Wortschatzmessung	(f)	mesure du vocabulaire		
41	Wortvertrautheitsliste	(f)	liste de familiarité		
42	Zusammensetzung	(f)	combinaison	(f)	
43	Zweitsprache	(f)	langue seconde	(f)	second language
44			mot (m) courant		running word
45			vocabulaire thématique		thematic/topical vocabulary
46			<<mot d'environnement>>		environmental word (Michea)
47			mot de milieu [46]		
48			mot de structure [11]		
49			nom courant [4]		

APPENDICE D

CALCULS ET NOTES SUR LE CRITERE DE L'USAGE
(calcul de l'indice d'usage des mots A et B)

0. DONNEES

A - 1 1 1 1 1 1 204

B - 30 30 30 30 30 30 30

A et B ont la même fréquence F = 210

A et B ont la même répartition r = 7

$\bar{x} = \Sigma x_i/n$ dans le cas présent 210/7 = 30

I. CALCUL INDICE DE DISPERSION: (Cet indice tient compte du nombre de sous-ensembles où le mot apparaît et des sous-fréquences comparées à la fréquence théorique)

a) Calcul de la variance: (Elle est la moyenne des carrés des déviations par rapport à la fréquence théorique)

	A			B		
	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
x_1	1	-29	841	30	0	0
x_2	1	-29	841	30	0	0
x_3	1	-29	841	30	0	0
x_4	1	-29	841	30	0	0
x_5	1	-29	841	30	0	0
x_6	1	-29	841	30	0	0
x_7	204	+174	30276	30	0	0
		0	35322		0	0

$\sigma^2A = 35322/7 = 5046$ $\sigma^2B = 0/7 = 0$

b) Calcul de l'écart-type: (L'écart-type est la racine carrée de la variance)

$\sigma A = \sqrt{5046} = 71.04$ $\sigma B = 0$

c) Calcul du coefficient de variation

$V = \frac{\sigma}{x}$; $V_A = 71.04 = 2.368$ $V_B = 0$

d) Indice de dispersion

$$D = 1 - \frac{V}{\sqrt{n - 1}}$$

$$D_A = 1 - \frac{2.368}{\sqrt{7 - 1}} = 1 - \frac{2.368}{2.449} = 0.04$$

$D_B = 1 - 0 = 1$

II. L'INDICE D'USAGE (U) = F X D

$U_A = F \times D_A = 210 \times 0.04 = 8.4$

$U_B = F \times D_B = 210 \times 1 = 210$

Nous constatons que l'usage équivaut à la fréquence mais seulement dans le cas où la fréquence est également répartie entre les sous-ensembles, c'est-à-dire quand D = 1.

APPENDICE E

INDEX DES AUTEURS CITES DANS LE CORPS DU TEXTE
(Les chiffres renvoient aux numéros des pages)

ARDOUIN	15
AUBER	74
BAKONYI, H.	9, 33-35, 39, 73, 75
BLOOMFIELD	61
BUCHANAN	18
CHATELANAT	48
COLEMAN	9, 30, 31
DROESCHER	101
DROTTENS	17
DUBOIS	17
EATON	28
ENGEL, E.F.	30, 32
ERK, H.	101
FRAISSE	16
GARCIA HOZ	13, 18
GOETHE	53
GOUGENHEIM	13, 17
GRIESBACH, H.	44
GUIRAUD	5
HARTMAN	4, 9
HAUCH	9, 31, 33, 41-43, 73

HENMON	28
HOHLFELD	35
HORN	11
JUILLAND	4, 13, 14
KAEDING	3, 8, 9, 12, 20, 23-28, 29, 30, 32, 34, 39, 40, 49, 57, 63, 72, 73, 75
KOEKKOEK	101
KEIL, M.	56-57, 72, 73, 75
KEIL, R.D.	4, 5, 9, 10
KELLER	5
KENISTON	13, 19
KOPPS, H.	54-56
LADO, R.	101
LAFORGE	5, 9, 12, 16, 17
LANG	48
LEIBNIZ	9
LESSING	43
LORICHE, R.	50
MACKEY, W.F.	9, 12, 15, 16, 100
MAROUZEAU	4
MASSERENTI	17
MATTUTAT	40
MEIER, Helmut	10, 39, 63
MEIER, Hermann	30-31, 32
MELZER, F.	51
MEYER	8

MICHEA	4, 15, 20, 39, 44-47, 101
MORGAN	9, 28-30, 32, 35, 36, 37, 73
MÜLLER, Ch.	13
MÜNSTER, T.	54-56
NASH	4
NOSTRAND	9
OEHLER, H.	39
OGDEN, C.K.	16
PFEFFER	3, 8, 39, 57-63, 72, 73, 75
PRICE	5
PURIN	31-32, 35, 37, 38
RICHARDS	9, 12, 17, 18, 20, 72
SAVARD	9, 12, 15, 16, 18, 20, 72
SCHILLER	43
SCHINNERER	31
SCHMIDT, P.	53
SCHNEIDER	9, 74
SCHULTZ, H.	33
SCHULTZE	39
SCHULZ, D.	44
SIMMONS	43
SLABY, H.	51
STEGER, H.	56-57, 72, 73, 75
THORNDIKE	11, 13
TOLLENAERE de	8, 9
UHLIG	48

ULLMANN	61
VAIL, C.	32-33, 35, 37, 38, 73
VANDER, Beke	13
WADEPUHL	30, 31, 35, 37, 38, 41, 73
WÄNGLER, H.	39, 48-50, 73
WENDT	31
WOLF, F.	51-53

APPENDICE F

RESUME
DU TRAVAIL

Le présent travail fait un inventaire historique des travaux de lexicométrie allemande de 1898 à 1971. Presque la moitié de ces travaux sont analysés d'après une grille qui permet de faire ressortir, entre autres, les critères de sélection et les méthodes de travail utilisés par les chercheurs germanistes.

Il est démontré qu'il existe de nombreuses listes de mots pour la seule langue allemande (une soixantaine). Elles ont été dressées selon les critères de fréquence, répartition, disponibilité et empirisme; le premier et le dernier de ces critères étant les plus employés.

Le terme lexicométrie est proposé en 1965 par Rolf Dietrich Keil. Bakonyi suggère une étude spéciale de la disponibilité auprès des enfants et des sourds-muets, l'étude des mots usuels. Steger et Keil ajoutent avec leur méthode de *Wortkonstellation* un nouvel aspect à la disponibilité: il s'agit de trouver des mots plus aptes à mener l'étudiant à la conversation. C'est Kaeding qui donne l'exemple d'une recherche scientifique en lexicométrie. Parmi les contemporains Pfeffer est sans doute le plus remarquable notamment pour l'allemand de nos jours.

Le travail se termine sur un tableau des langues du monde traitées ou en traitement lexicométrique, un tableau des différents centres de traitement et un petit lexique des termes lexicométriques.

Dans la même série:

B-1 *L'utilisation de l'ordinateur en lexicométrie.*
Savard, Jean-Guy

B-2 *L'ordinateur et l'analyse grammaticale.*
Mepham, Michael

B-3 *Concept Categories as Measures of Culture Distance.*
Mackey, William F.

B-4 *L'université bilingue.*
Verdoodt, Albert

B-5 *La rentabilité des mini-langues.*
Mackey, William F.

B-6 *The Computer in Automated Language Teaching.*
Mackey, William F.

B-7 *The Three-Fold Objective of the Language Reform in Mainland China in the Last Two Decades.*
Chiu, Rosaline Kwan-Wai

B-8 *Un test télévisé.*
Savard, Jean-Guy

B-9 *Sociolinguistic History, Sociolinguistic Geography, and Bilingualism.*
Afendras, Evangelos A.

B-10 *Mathematical Models for Balkan Phonological Convergence.*
Afendras, Evangelos A.

B-11 *Stability of a Bilingual Situation and Aromanian Bilingualism.*
Afendras, Evangelos A.

B-12 *More on Informational Entropy, Redundancy and Sound Change.*
Afendras, Evangelos A. and TZANNES, Nicolas S.

B-13 *Relations entre anglophones et francophones dans les syndicats québécois.*
Verdoodt, Albert

B-14 *Multilingual Communication in Nigeria.*
Iso, Asi Otu and Afendras, Evangelos A.

B-15 *The Language Factor in Maori Schooling.*
Richards, Jack C.

B-16 *Diffusion Processes in Language Prediction and Planning.*
Afendras, Evangelos A.

B-17 *A Non-Contrastive Approach to Error Analysis.*
Richards, Jack C.

B-18 *Research Possibilities on Group Bilingualism: A Report.*
Kloss, Heinz.

B-19 *Interference, Integration and the Synchronic Fallacy.*
Mackey, William F.

B-20 *A Psycholinguistic Measure of Vocabulary Selection.*
Richards, Jack C.

B-21 *A Pilot Study on the Ability of Young Children and Adults to Identify and Reproduce Novel Speech Sounds.*
Afendras, Evangelos A., Yeni-Komshian, G. and Zubin, D.

B-22 *Can One Measure a Sprachbund? A Calculus of Phonemic Distribution for Language Contact.*
Afendras, Evangelos A.

B-23 *Stochastic Processes for Diachronic Linguistics.*
Afendras, Evangelos A. and Tzannes, Nicolas S.

B-24 *Structures ethniques et linguistiques au Burundi, pays "unimodal" typique.*
Verdoodt, Albert

B-25 *Error Analysis and Second Language Strategies.*
Richards, Jack C.

B-26 *Graduate Education in Foreign Languages.*
Mackey, William F.

B-27 *La question scolaire en Alsace: Statut Confessionnel et Bilinguisme.*
Kauffmann, Jean

B-28 *Polychronometry: the Study of Time Variables in Behavior.*
Mackey, William F.

B-29 *La diglossie au Québec: limites et tendances actuelles.*
Chantefort, Pierre

B-30 *Literary Biculturalism and the Thought-Language-Culture Relation.*
Mackey, William F.

B-31 *Chicago Conference on Child Language - Preprints.*
Centre international de recherches sur le bilinguisme (réd.)

B-32 *La distance interlinguistique.*
Mackey, William F.

B-33 *Options politiques fondamentales de l'Etat plurilingue.*
Plourde, Gaston

B-34 *Social Factors, Interlanguage and Language Learning.*
Richards, Jack

B-35 *Analyse des erreurs et grammaire générative: la syntaxe de l'interrogation en français.*
Py, Bernard

B-36 *Anglicization in Quebec City.*
Edwards, Vivian

Autres publications du C.I.R.B.

Série A - Etudes

A-1 SAVARD, Jean-Guy et RICHARDS, Jack. *Les indices d'utilité du vocabulaire fondamental français.* Québec, Les Presses de l'Université Laval, 1970. 169 p.

A-2 KLOSS, Heinz. *Les droits linguistiques des Franco-Américains aux Etats-Unis.* Québec, Les Presses de l'Université Laval, 1971. 80 p.

A-3 FALCH, Jean. *Contribution à l'étude du statut des langues en Europe.* Québec, Les Presses de l'Université Laval, 1973. 300 p.

A-4 DORION, Henri et MORISSONNEAU, Christian (éditeurs). *Les noms de lieux et le contact des langues / Place Names and Language Contact.* Québec, Les Presses de l'Université Laval, 1972. 374 p.

A-5 LAFORGE, Lorne. *La sélection en didactique analytique.* Québec, Les Presses de l'Université Laval, 1972. 675 p.

(sous presse)

A-6 TOURET, Bernard. *L'aménagement constitutionnel des Etats de peuplement composite.* Québec, Les Presses de l'Université Laval, (1973). 256 p.

A-7 MEPHAM, Michael. *Computation in Language Text Analysis.* Québec, Les Presses de l'Université Laval, (1973).

Série C - Publications extérieures

C-1 SAVARD, Jean-Guy. *La valence lexicale.* Paris, Didier, 1970. 236 p.

C-2 MACKEY, William F. *Bilingualism as a World Problem / Le bilinguisme phénomène mondial.* Montréal, Harvest House, 1967. 119 p.

C-3 MACKEY, William F., SAVARD, Jean-Guy et ARDOUIN, Pierre. *Le vocabulaire disponible du français.* Paris, Didier, 1971. 2 volumes, 900 p.

C-4 STERN, H.H. (réd.) *L'enseignement des langues et l'écolier.* Rapport d'un colloque international, traduit par le CIRB sous la direction de William F. Mackey, Institut de l'Unesco pour l'éducation, Hambourg, 1971. 254 p.

C-5 KLOSS, Heinz. *Laws and Legal Documents Relating to Problems of Bilingual Education in the United States.* Washington, D.C., Center for Applied Linguistics, 1971. 92 p.

C-6 MACKEY, William F. *Principes de didactique analytique.* Traduction française *(Language Teaching Analysis)* et mise à jour de Lorne Laforge. Paris, Didier, 1972. 713 p.

(sous presse)

MACKEY, William F. et VERDOODT, Albert (éditeurs). *The Multi-National Society.* Rowley (Mass.), Newbury House.

Collection Studies in Bilingual Education (Newbury House, Rowley, Mass.)
W.F. Mackey - General Editor

C-100 MACKEY, William F. *Bilingual Education in a Binational School: a study of equal language maintenance through free alternation.* 1972. 185 p.

C-101 SPOLSKY, Bernard. *The Language Education of Minority Children: selected readings.* 1972. 200 p.

C-102 LAMBERT, Wallace E. et TUCKER, G. Richard. *Bilingual Education of Children: The St. Lambert Experiment.* 1972. 248 p.

(sous presse)

ANDERSSON, Theodore et MACKEY, William F. *Bilingualism in Early Childhood.*

KLOSS, Heinz. *The American Bilingual Tradition in Education and Public Administration.*

GAARDER, A. Bruce. *Essays on Bilingual Schooling in the United States: Its Present Role and Potential Scope.*

COHEN, Andrew. *A Sociolinguistic Approach to Bilingual Education.*

Série E - Inventaires

(en préparation)

KLOSS, Heinz, directeur de l'édition; MCCONNEL, G., directeur adjoint de l'édition. *Composition linguistique des Nations du monde.* Vol. 1: *Les Indes et le Proche-Orient.*

KLOSS, Heinz, directeur de l'édition; MCCONNEL, G., directeur adjoint de l'édition. *Les langues alphabétisées du monde.* Vol. 1: *Les Amériques.*

Série F - Bibliographies

F-1 SAVARD, Jean-Guy. *Bibliographie analytique de tests de langue / Analytical Bibliography of Language Tests.* Québec, Les Presses de l'Université Laval, 1969. 372 p.

F-2 CHIU, Rosaline Kwan-wai. *Language Contact and Language Planning in China (1900-1967). A Selected Bibliography.* Québec, Les Presses de l'Université Laval, 1970. 273 p.

F-3 MACKEY, William F. (rédacteur). *Bibliographie internationale sur le bilinguisme / International Bibliography on Bilingualism.* Les Presses de l'Université Laval, 1972. 757 p.